OSCA, an Optimised Stellar Coronagraph for Adaptive optics

Samantha Jayne Thompson

Thesis submitted for the Degree of Doctor of Philosophy
of the University of London

Department of Physics & Astronomy
UNIVERSITY COLLEGE LONDON

September 2004

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

ABSTRACT

Described here is the design, manufacturing, testing and commissioning of a coronagraph facility for the 4.2 metre William Herschel Telescope (WHT) and its adaptive optics system (NAOMI). The use of the NAOMI adaptive optics system gives an improved image resolution of $\sim$ 0.15 arcseconds at a wavelength of 2.2μm. This enables the Optimised Stellar Coronagraph for Adaptive optics (OSCA) to null stellar light with smaller occulting masks and thus allows regions closer to bright astronomical objects to be imaged.

OSCA is a fully deployable instrument which when in use leaves the focus of the NAOMI beam unchanged. This enables OSCA to be used in conjunction with a number of instruments that have already been commissioned at the WHT. The main imaging camera used with OSCA is INGRID; a 1024 $\times$ 1024 pixel HgCdTe cooled short-wave infra-red (SWIR) detector at the NAOMI focus. OSCA also has the option of being used in conjunction with an integral field spectrograph for imaging at visible wavelengths. OSCA provides a selection of 10 different occulting masks with sizes of 0.25 - 2.0 arcseconds in diameter, including two with novel gaussian profiles. There is also a choice of two different sized Lyot stops (pupil plane masks). A dichroic placed before the AO system can give improved suppression performance when occulting masks larger than the seeing disk are used.

Also presented are results from observing time with the OSCA system, which highlight the challenges faced by astronomers to obtain high contrast, high resolution images from ground based telescopes. At a time during which there is much activity towards terrestrial planet finding, questions as to the system requirements required for such a task are discussed.

To Mum and Dad,

for always believing.

An artist's impression of a possible planetary system around the star Fomalhaut. One of the major driving forces in the development of high contrast imaging systems in astronomy is for the study of proto-planetary disks with the eventual aim to directly image exo-planets.

Picture courtesy of David A Hardy.

ACKNOWLEDGEMENTS

There are numerous people I feel I must thank for their help and advice over the course of my PhD. Firstly my supervisor Peter Doel, for his help and guidance and patient reading over of this thesis. Thanks to the past and present members of the Optical Science Laboratory who have provided support along the way, especially David Brooks for much helpful advice and use of his secret tool stash, without which OSCA could not have been assembled! My gratitude also extends to the many astronomers at UCL and elsewhere who responded to my request for interesting commissioning targets and explained their astronomical significance to me. My work has been greatly enhanced and inspired by my dabblings into these different fields.

I must also thank the many people that worked on the OSCA project with me: Andy Charalambous for explaining engineering matters and tricks of AutoCAD, Richard Bingham for optical design deliberations and the many people at Durham University, in particular Nirmal Bissonauth who came on the first OSCA commissioning and an extra thanks to Paul Clark for the best soldering lesson I've ever had! I would also like to thank the staff at the William Herschel Telescope for their support during commissioning and observing, particularly Sebastian Els as the OSCA support scientist.

Thanks to Sam Searle, Dugan Witherick and Jay Abbott, my office mates, for their friendship and listening to my rantings! To my other good friends who I've neglected over the past few months to finish this - thanks for putting up with me.

Last, but by no means least, a special thanks to Richard, his love and support has seen me through this PhD, enduring my ramblings on obscure ideas and my research papers all over our house; ever the marvellous chef and MATLAB guru.

"Science is Fun!"

–Mr. Jennings (school physics teacher)

"I am among those who think that science has great beauty. A scientist in his laboratory is not only a technician: he is also a child placed before natural phenomena which impress him like a fairy tale."

–Marie Curie

Contents

List of Figures

List of Tables

PREFACE

There are many people involved in the development of a facility class instrument such as OSCA. I have included the work of these people in this thesis for completeness and will now mention who they are and the key roles they had in the OSCA project. This is also to clarify the work which I, the only student on the project contributed. The list of people involved in the OSCA project and their roles is as follows:

Peter Doel (APD)	Principle investigator	University College London (UCL)
Samantha Thompson (SJT)	Student of APD	UCL
Richard Bingham (RGB)	Optical design	UCL
Andy Charalambous (AC)	Mechanical design	UCL
Richard Myers (RMM)	Co-I + NAOMI	University of Durham (DUR)
Gordon Love	Co-I	DUR
Paul Clark (PC)	Electronic design	DUR
Nirmal Bissonauth (NB)	EPICS control software	DUR
Nigel Dipper	Control software	DUR
Gordon Talbot (RGT)	Engineering	Isaac Newton Group (ING)
Craige Bevil (CB)	GUI control	ING
Paul Jolley (PJ)	WHT engineering	ING
Tom Gregory (TG)	Optics	ING

My initial role was in the modelling of the NAOMI/OSCA system, to determine and constrain the important system parameters. The information gained from this was then

used in discussion with the optical and mechanical designers so that we could arrive at a design for the coronagraph with acceptable error margins, to ensure these would not compromise the performance targets. In the mechanical case this was calculating the acceptable tolerances in positioning of the system, since OSCA was to be auto-deployable in and out of the NAOMI beam and there would be a rotating wheel for occulting mask selection. Such numbers enabled AC to pick the most suitable components. I also sourced many of the commercially available components that were used in OSCA - these included the Newport SR50 rotation stage, mirror mounts and stepper motor. Although AC was responsible for most of the drawing and mechanical design, I designed an anti-scatter mask for the focal substrates, the mirror clips and spacers to modify the commercial mounts and I designed the plate holder used in the photolithography production of the focal masks.

For the optical case RGB was responsible for the overall optical design, and an initial design had already been completed before my involvement in the project had started. The design has undergone two revisions since then. I was responsible for the design and method of fabrication of both the occulting masks and the pupil plane masks. I also decided upon the best optical materials available for the mirrors and windows, and oversaw the manufacturing of the custom components (windows and off-axis paraboloids).

Initial assembly, alignment and testing of the opto-mechanical system and the OSCA-side electrical wiring was performed largely by myself. I also assisted the team from Durham with the electronic systems integration with the OSCA opto-mechanics and subsequent testing of the control system. First commissioning at the telescope was undertaken by APD and myself (optics/mechanics), NB (electronics/software) and ING staff. An upgrade and additional testing were also performed at a later date by APD and myself. Lastly, I took part in an observing run using the NAOMI-OSCA system.

This thesis is organised as follows:

Chapter 1 - brief introduction to subjects involved

Chapter 2 - the design and manufacturing of the components in OSCA

Chapter 3 - computer simulations of the system

Chapter 4 - construction, testing, commissioning of OSCA and on-sky results

Chapter 5 - discussion

Appendix A - additional photographs taken throughout the OSCA project

Appendix B - additional CAD drawings

CHAPTER 1

Introduction

The study of coronagraphy and high contrast imaging requires an understanding of the theory of Fourier optics. A coronagraph is essentially a spatial filter with a filter mask at the telescope focus and at the subsequent pupil plane. The concepts of this theory are introduced here and following on from this is a discussion on the principles of coronagraphic imaging. An overview of adaptive optics and the system at the William Herschel Telescope is also given.

1.1 Fourier Optics

The Fourier transform is widely used in the analysis of physical systems. In optics the relationship between the two dimensional (2-D) Fourier transform and the various phenomena such as diffraction, interference etc. is especially important. In §3 Fast Fourier Transforms (FFTs) have been used in a computer model of the OSCA coronagraphic system. The diffraction pattern produced when light passes through an aperture is related to the Fourier transform of the aperture function. Since diffraction is an intermediate step in image formation a diffraction pattern can be manipulated to alter certain factors in the final image.

To justify the use of the Fourier transform in the modelling of an optical system consider Figure 1.2. Starting from Huygens principle, each element of area $\mathrm{d}A$ of a given aperture

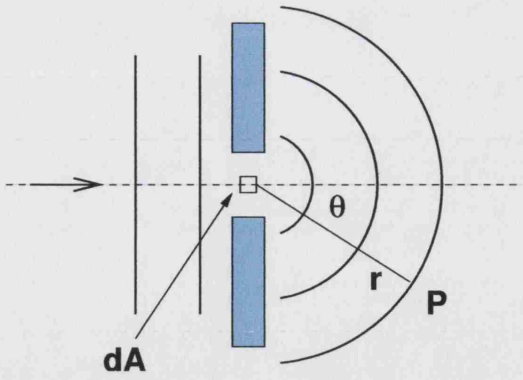

Figure 1.1: Diagram showing a Huygens-Fresnel wavelet originating from a region $\mathrm{d}A$, in an aperture of area A

can be considered as a point source and a source of spherical wavelets (Figure 1.1). At a point P, a distance r from the aperture, the wave contribution $\mathrm{d}\psi$ is given by the Fresnel wavelet equation:

$$\mathrm{d}\psi = \frac{a(x,y)}{r}\, e^{i(\omega t - kr)}\, \mathrm{d}A \tag{1.1}$$

Where $a(x,y)$ is the amplitude per unit area of a light wave and k is the wave number, defined as usual to be $\frac{2\pi}{\lambda}$. Note, the coordinates in the aperture plane are x, y and in the plane containing the point P are u, v.

As $r \to \infty$ the wavefronts can be taken to be planar and the wavelets take the form of the Fraunhofer approximation. It is preferable to work in the Fraunhofer regime since the calculations are simplified compared to the Fresnel (or near-field) approximation. The condition for Fraunhofer diffraction can be shown to be:

$R \gg \frac{k(x^2+y^2)_{\max}}{2}$ (Goodman 1996, ch.4)

where R is the distance of the point P (Figure 1.2) from the aperture origin and x and y are the largest dimension in the aperture; even for small apertures this makes the Fraunhofer distance very large. However, the Fraunhofer diffraction pattern can be observed much closer than this if a suitable lens is positioned between the observer and the aperture (i.e. apparent R is very large) or if the aperture is illuminated by a spherical wave converging toward the observer. Then we can take $r \simeq R$ for all wavelets, and the $1/r$ term in Equation 1.1 becomes a constant. The equation can then be re-written as:

$$\mathrm{d}\psi = \frac{a(x,y)}{R}\, e^{i\omega t}.\, e^{-ikr}\, \mathrm{d}A \tag{1.2}$$

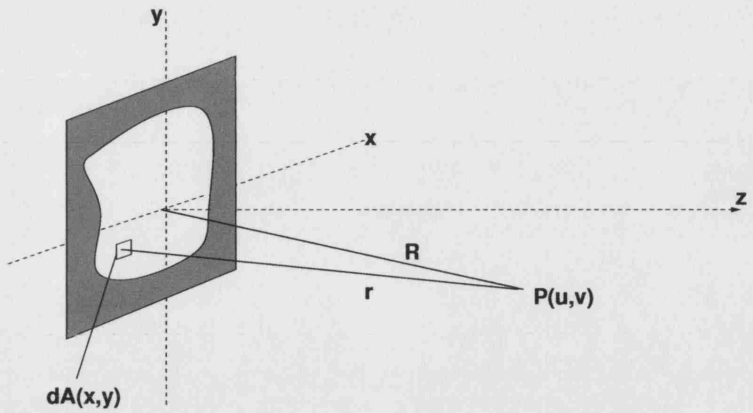

Figure 1.2: Diagrammatic representation of an aperture in the xy-plane (object plane), wave travel in the +z direction and diffraction pattern in the uv-plane (transform plane).

Solving for r in terms of x, y and u, v (Wilson 1995, ch.5), yields the expression:

$$\mathrm{d}\psi = \frac{a(x,y)}{R}\, e^{i(\omega t - kR)} . e^{ik(ux+vy)/R}\, \mathrm{d}A \tag{1.3}$$

By the principle of superposition, the net effect of diffraction is obtained by summing the amplitudes from all the spherical wavelets over the entire aperture. Usually only the relative intensity values are required, so the terms in Equation 1.3 that are independent of x and y can be set to 1. This summation can be expressed as the integral:

$$\psi(u,v) = \int\int_{-\infty}^{+\infty} A(x,y)\, e^{ik(ux+vy)/R}\, \mathrm{d}x\, \mathrm{d}y \tag{1.4}$$

$A(x,y)$ is the aperture function, so that:
$A = 1 \Rightarrow$ transparent
$A = 0 \Rightarrow$ opaque
$0 < A < 1 \Rightarrow$ partially opaque
$A = A\, e^{i\alpha} \Rightarrow$ phase shift α

Since the observable effect is the intensity distribution this is given by $|\psi|^2$. The general form for the 2-d Fourier transform is:

$$F(u,v) = \int\int_{-\infty}^{\infty} f(x,y)\, e^{-2\pi i(ux+vy)} \mathrm{d}x\, \mathrm{d}y \tag{1.5}$$

So comparing Equation 1.4 with 1.5 it can be seen that the diffraction pattern is the Fourier transform of the aperture function and conversely the inverse Fourier transform of the diffraction pattern will produce the aperture function.

Fourier transforms are well known for their use in signal processing, where the transform is between time and frequency; in image processing and optics the corresponding analogy is between *space* and *spatial frequency*. This translates to a feature of size d in the aperture plane being distributed over $1/d$ in the transform plane, i.e. small features (detail, sharp edges etc.) are said to be 'high frequency', where as the larger scale structures are often described as 'low frequency'. Thus the Fraunhofer diffraction pattern $\psi(u, v)$, also contains information about the spatial frequency spectrum of $A(x, y)$.

Computation of diffraction patterns generally uses FFTs which are discrete algorithms for the Fourier transform (expressed in its continuous form in Eq.1.5). The size of the matrix used to represent an aperture will place a limit on the number of spatial frequencies sampled and hence the level of detail that will be present in the transformed image. Again this is discussed further in §3 when choosing a suitable sized matrix for the OSCA simulations.

1.2 Coronagraphy

The first coronagraphs (Lyot 1939), as their name suggests, were created to study the solar corona in the absence of an eclipse. The idea was expanded to use the system on stellar objects, to suppress bright starlight and thus see fainter objects closer to the star. These Lyot coronagraphs essentially consist of an opaque spot placed at the focus of the telescope and an undersized mask of the aperture in the pupil plane after the focus. Figure 1.3 is a simple diagram illustrating this layout.

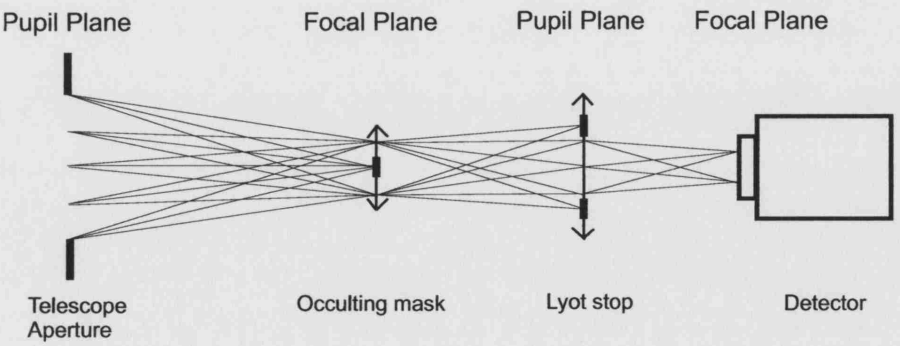

Figure 1.3: General optical schematic of a coronagraph (Malbet 1996)

The function of these components is to first block most of the light from the central bright star and secondly to remove the remaining diffracted light from the telescope aperture. A complete description can be given using the Fourier theory of optics. This theory uses Fourier transforms to describe the change from focal plane to pupil plane, as has been summarised in §1.1. So in a coronagraphic system the focal plane can be represented by the Fourier transform of the telescope aperture function. Thus by inserting a stop at the focus unwanted starlight is removed along with the low spatial frequency components from the aperture function. When the inverse transform is performed and the pupil plane viewed, the majority of the light remaining is associated with the edges of the aperture which corresponds to high spatial frequencies in the focal plane. This is demonstrated in Figure 1.4. The representative telescope aperture was illuminated in the laboratory with a broad collimated laser beam (613nm) and the emerging light brought to a focus. A small (300μm) opaque disk was positioned at this focus and the light brought back to a collimated beam and imaged on an SBIG CCD camera. As the figure shows, the remaining light at this re-imaged aperture is concentrated around the edges. By inserting an undersized aperture in this plane (a Lyot stop), this excess light can easily be removed which otherwise would produce bright diffraction effects in a subsequent focus. For an off-axis object (beyond the influence of the focal mask), the distribution of light in the pupil is unaltered, so the Lyot stop will preferentially suppress the light from the central star. Naturally the Lyot stop itself will introduce unwanted diffraction effects, and under-sizing it too much will take out much of the spatial frequency information that makes up the focal image. So care must be taken to get the balance right.

When attempting to observe faint stellar objects a long integration time on the source is preferred to obtain a reasonable signal to noise ratio. However, if a bright object is in the observed vicinity and a coronagraph is not used to remove the unwanted light then an observer would encounter problems such as CCD saturation (even at very short exposures) and lower signal to noise ratio, due largely to the increase in read-out noise (more frames required for equivalent integration time), spurious reflections, scattered light and bright diffraction spikes from the vanes in the telescope aperture that reduce image contrast.

Improvements in coronagraphy are driven by the requirements to observe closer to bright astronomical objects and with greater suppression of their light than has been done before. Early stellar coronagraphs were limited by atmospheric seeing conditions and so the

occulting mask sizes had to be large to allow for this movement. Since the 1980's astronomers have been very interested in a young, nearby, star called Beta Pictoris. Figure 1.5 is one of the first coronagraphic images of β Pic and reveals an extended emission either side of the star which is due to a disk of dust surrounding the star which we view nearly edge-on.

The coronagraph used to obtain these data has since been de-commissioned but there are a number of other coronagraphs available for general use today, which are summarised in Table 1.1. The β Pic disk has been observed in increasing detail by some of these instruments in a search for evidence of planetary formation. The requirements are to observe the disk as close in to the star as possible without becoming overwhelmed by the direct starlight. To build smaller effective occulting masks requires the star PSF to be near diffraction-limited and stable (does not move around). This has been achieved using the

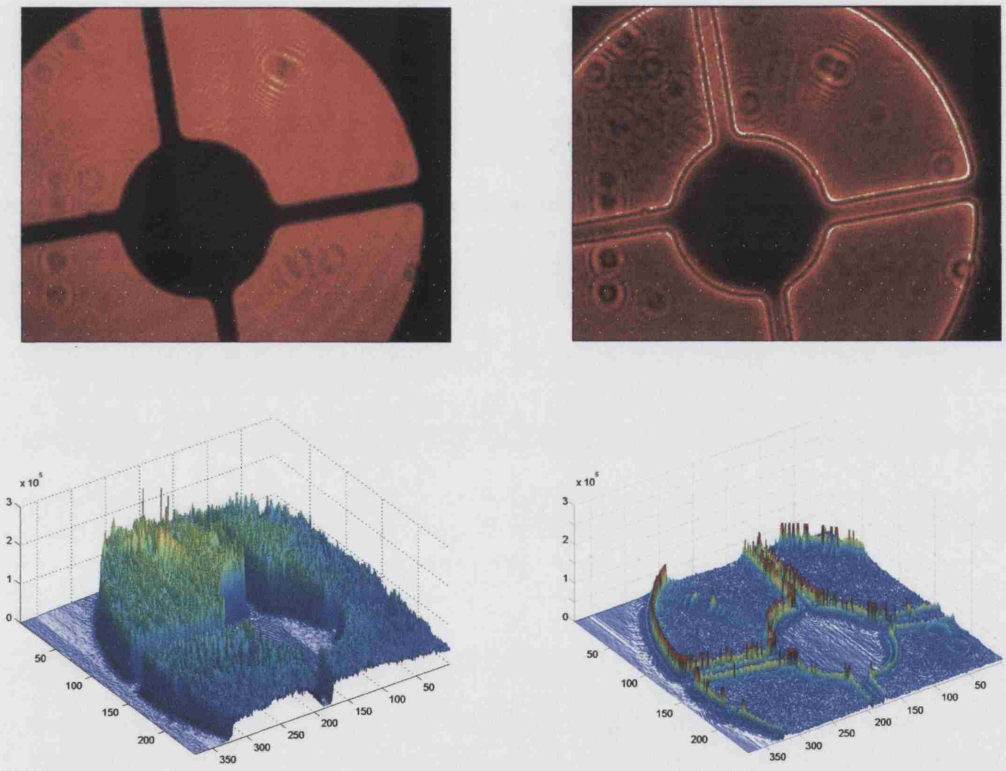

Figure 1.4: Laboratory images showing the effect in the pupil plane of placing a small occulting spot in the focus up-beam. On the left is the pupil without a focal-spot and underneath a graph displaying the light intensity in the image. On the right is the same pupil but with a 300μm opaque disk at the focus.

Table 1.1: Summary of current coronagraphs[1] and facilities. Occulting mask sizes are diameters in arcsec and the type is denoted by: a - hard edged on silica, b - hard edged on wires, c - gaussian HEBS mask (full greyscale), for visible only, d - set of tapered bars on single plate, cool - indicates masks are in a cryo-chamber.

Coronagraph	Telescope	Occulting masks	FOV (arcsec)	Pixel scale (mas)	AO Correction	λ (μm)	Latitude (degrees)	Other features
STIS	HST - 2.4m	0.6-3.0 d	52	50.7	n/a	0.58 peak (white light)	space	
NICMOS	HST - 2.4m	0.6	19.2	75	n/a	0.8 - 2.5	space	
ACS	HST - 2.4m	1.8, 3.6 a	26	25	n/a	0.2 - 1.1	space	
NIRC2	Keck2 - 10m	0.1-2.0 a cool	10, 20, 40	10, 20, 40	TTM, 349 act.	1 - 5	+19 49.6	
CONICA	VLT, Yepun - 8.2m	0.7 a,b; 1.4 b cool	14 - 56	13.3-54.7	TTM, 185 PZT	1 - 5	-24 37.5	
CIAO	Subaru - 8.2m	0.1-4.0 a cool	11.8, 22.3	11.5, 21.7	TTM, bimorph	1 - 5	+19 49.6	linear polarimetry
PHARO	Palomar - 5m	0.41, 0.91 cool	25 - 40	25-40	TTM, 349/241 act.	1.2 - 2.3	+33 21.4	
OSCA	WHT - 4.2m	0.3-2.0 a; 0.5 c	20	38	TTM, 76 segmented	0.4 - 2.4	+28 45.5	OASIS - IFU
CASPIR	SSO, ANU - 2.3m	2.0, 5.0 cool	126	500	tip/tilt secondary	1 - 5	-31 16.6	

[a]This refers to facility class/permanently installed coronagraphs as of January 2004. There may be other research grade coronagraphs in operation that are not listed here.

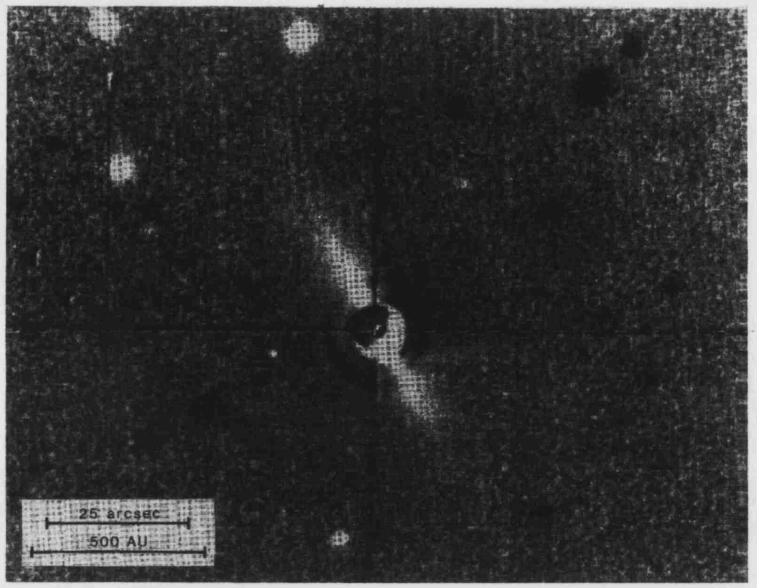

Figure 1.5: First coronagraphic image of the disk around β Pictoris (Smith & Terrile 1984) obtained by the 2.5 m telescope at Las Campanas Observatory in Chile. The occulting mask used to obtain this image was 7 arcsec in diameter and has been divided by an image of α Pictoris to reveal the edge-on circumstellar disk, shown here to extend 25 arsec (400 AU) to the northeast and southwest of the star (North is up).

Hubble Space Telescope (HST) which removes the restriction of the Earth's atmosphere and to a lesser extent the development of adaptive optics systems which are discussed in §1.3. Figure 1.6 is a detailed image (0.1 arcsec resolution) of the β Pictoris disk taken with the HST and the Space Telescope Imaging Spectrograph (STIS). The occulting mask in this case is a wedge shape (tapered bar). The disk in these images shows a warping in its shape, a hypothesis for this phenomenon is that there are one or more planets in orbit about the star (Burrows *et al.* 1995).

Goals for the future are to use coronagraphs to suppress a star's light to such an extent that any planets in orbit around it will be directly detectable. Designs for such instruments are already in progress, e.g. the Jovian Planet Finder (JPF) (Clampin *et al.* 2001), a space based instrument planned for launch in 2007. The opening figure of this thesis is an illustration of these future ambitions. Numerous designs for coronagraphs have been suggested and simulated over the years and will not be described in any detail here unless

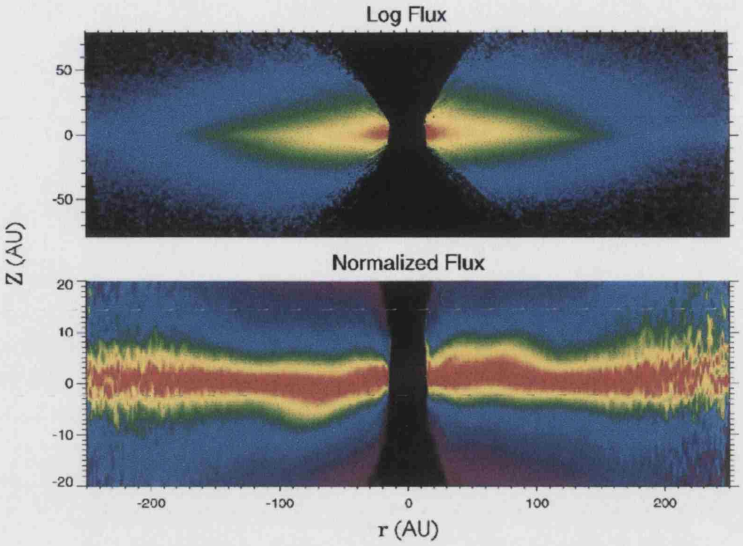

Figure 1.6: Hubble Space Telescope, STIS/CCD coronagraphic images of the β Pictoris disk. Half width of the occulted region is 0.75 arcsec. The top image shows the disk at a logarithmic stretch. The bottom images shows the disk normalised to the maximum flux and vertical scale expanded (Heap *et al.* 2000).

pertinent to the OSCA coronagraph. Some different designs have been investigated in §3. The most notable designs published over recent years deserve mentioning and include:

- Phase-mask or nulling stellar coronagraph (Roddier & Roddier 1997)
- The four-quadrant phase-mask coronagraph (Rouan *et al.* 2000)
- Achromatic interfero coronagraph (Baudoz *et al.* 2000)
- Various pupil masking techniques, generally referred to as 'apodization' (e.g. Watson *et al.* 1991; Nisenson & Papaliolios 2001; Aime *et al.* 2001; Debes *et al.* 2002)

However in practice many of these designs cannot as yet be manufactured to the specifications required or need near perfect wavefronts (only suitable for space based telescopes). So most coronagraphs in use on telescopes today are based on the original Lyot coronagraph design.

1.3 Adaptive Optics

With the advent of adaptive optics (AO) for telescopes (e.g. Rigaut *et al.* (1991)), random optical wavefront distortions can be corrected in real-time. This has enabled ground based telescopes to break free of the angular resolution constraints imposed by the turbulent atmosphere and work towards diffraction-limited imaging. To gain a more detailed understanding of the universe there is a requirement for increasingly larger telescopes with higher resolution capabilities. Contrary to popular belief, increasing the aperture of a terrestrial telescope beyond a certain diameter (local seeing conditions dependent) does not improve the resolution of its image, it merely allows more light to be gathered and hence fainter, *unresolved* sources to be detected. This limiting diameter is approximately equal to r_0, the Fried parameter, which is a standard measure of the optical strength of atmospheric turbulence and corresponds to a coherence length for the optical phase distortions. As a function of both the local atmospheric conditions and wavelength, r_0 can be very variable over short time scales. Thus in the absence of any wavefront correction the angular resolution of the telescope is limited to $\sim \lambda/r_0$ instead of the diffraction limit $1.22\lambda/D$, (where D is the diameter of the telescope aperture). At a wavelength of 500nm, r_0 can vary from about 5cm (bad seeing) to 20cm (good seeing), so a telescope whose aperture is less than r_0 will have negligible loss of resolution due to the atmosphere. On the other hand for an aperture of size nr_0, where $n > 1$, as n gets larger the performance loss compared to the optimum increases and the need for adaptive optics becomes apparent.

The Strehl ratio S, is a good measure of optical quality and is used when measuring the performance of an AO system. It is defined as the ratio of the peak intensity of the PSF by an (aberrated) optical system to the corresponding intensity when that image is formed by an equivalent diffraction limited (aberration-free) system; so that $S = 1$ implies diffraction limited imaging. Under average seeing conditions on a non-AO telescope the estimated Strehl ratio in the infra-red (IR) is ~ 0.08 dropping to 0.01 in the visible. With AO it is expected that this can be improved to ~ 0.8 in IR and 0.2 - 0.3 in the visible (Hardy 1998). The loss of Strehl in AO systems can be attributed to the lack of correction for the high modes/frequencies in the wavefront; this produces a PSF with a more defined peak but with a broad base - the 'seeing halo'.

As turbulence in the atmosphere increases so does the width of the seeing disk and there

is a limit after which AO is unable to provide any correction. In reality most AO systems are not used if the atmospheric seeing conditions are worse than 1 arcsec.

1.3.1 Atmospheric correction methods

The current methods of wavefront correction involve the use of a wavefront sensor to detect the distortions and a deformable mirror to compensate for them. The most common wavefront sensor used is the Shack-Hartmann sensor, although curvature sensors are becoming more popular. Standard systems use just one wavefront sensor, those with more than one wavefront sensor are called multi-conjugate - such systems are being developed for the current generation of large telescopes, e.g. Gemini South (Ellerbroek *et al.* 2003) and VLT (Marchetti *et al.* 2004). There are a number of different types of adaptive mirror. These include a continuous faceplate mirror which is mounted on a bed of actuators which apply a force perpendicular to the mirror surface to deform it; a bimorph mirror which is a composite comprising layers of reflective coated glass, piezo-electric ceramic and support materials, an electric potential applied to electrodes on the back of the ceramic alters the local bending moment via the indirect piezo-electric effect and so alters the mirror surface shape; or a segmented mirror which consists of many smaller (usually square) mirrors mounted on actuators to adjust their tip, tilt and piston.

A wavefront sensor on a given telescope will have a minimum magnitude star it needs to sense from and for accurate correction the guide star should ideally be within the isoplanatic angle of the object being observed, a separation of 10-20 arcsec is usually adequate for infra-red observations. The isoplanatic angle is used to describe the angular size on sky over which wavefront distortions are correlated. Atmospheric tubulence effects are stronger for shorter wavelengths and so the isoplantic angle is smaller, it increases linearly with wavelength allowing greater sky coverage and better correction for AO at IR wavelengths. The wavefront used for correction may come from a natural guide star (NGS) or an artificially created laser guide star (LGS). A LGS is needed to improve sky coverage, particularly at visible wavelengths where the number of of suitable bright guide stars within the isoplanatic angle of the target is extremely low. When planning an adaptive optics observing program the criteria for NGS can sometimes be very difficult to meet, as there is only a small percentage of objects in the sky that have a suitable guide star in the near vicinity.

1.3.2 AO at the William Herschel Telescope

The adaptive optics system at the WHT, NAOMI (Nasmyth Adaptive Optics for Multi-purpose Instrumentation), consists of a single Shack-Hartmann wavefront sensor and a 76-element segmented mirror (see Figure 1.7). To accommodate a range of different seeing conditions and guide star magnitudes the NAOMI wavefront sensor has a number of different configurations, accommodated by a selection of different lenslet arrays in a wheel prior to the wavefront sensor camera. There are plans for a future upgrade to a Rayleigh laser guide star facility.

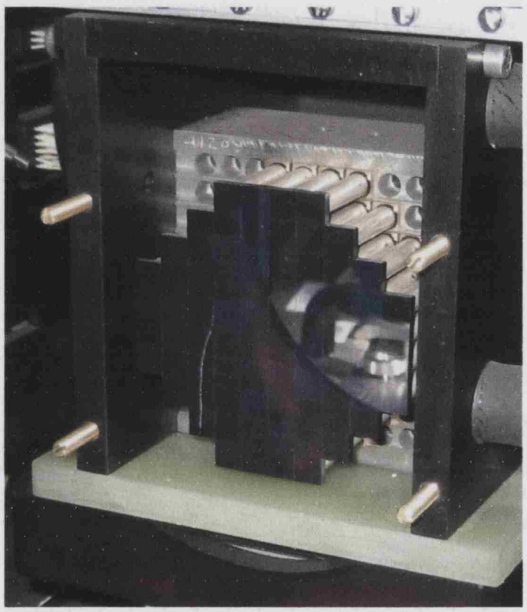

Figure 1.7: The deformable mirror in NAOMI consists of a 76-segment Thermotrex mirror

NAOMI has been at the telescope since 2000 and during early 2003 was moved to a new environment-controlled laboratory at the opposite Nasmyth platform. This should provide a more dust-free and thermally stable environment and so improve the performance of the system. In the current configuration the telescope pupil has a diameter on the deformable mirror of 7.3 segments (Myers *et al.* 2003), where each segment is 7.6 mm across. The centre four segments of the DM are not used since these are mostly obscured by the secondary mirror of the telescope, so the actual number of segments used for correction is 72. Since each segment is controlled by 3 actuators to give piston, tip and tilt, the NAOMI

mirror has 216 degrees of freedom. This is a reasonably high-order AO system and can offer partial correction at visible wavelengths. However, since the outer 24 segments of the DM lie mainly outside of the telescope pupil, most of the AO correction is made using only 48 segments (144 degrees of freedom).

The imaging camera used with NAOMI and OSCA is INGRID, a 1024 × 1024 element HgCdTe cryogenic cooled near-IR detector. The pixel scale when used with NAOMI is $\sim$ 0.04 arcsec/pixel. Prior to the detector but within the cryo-chamber are a set of 3 wheels containing broadband filters (Z - K), narrowband filters and a choice of pupil stops respectively.

1.4 High Contrast Imaging in Astronomy

Stellar coronagraphy is ideal to couple with an adaptive optics system; the bright central object that the coronagraph is trying to suppress can be used for wavefront sensing, removing the need to search for another suitable star to guide on. The combination of a coronagraph with an adaptive optics system presents the possibility of several interesting science studies. Compared to a standard coronagraphic imaging system, one with adaptive optics gives much higher spatial resolution and a high dynamic range allowing the environments of bright objects to be studied closer in than ever before. One of the first generation adaptive optics coronagraphs was the AOC, owned by John Hopkins University. It has been used successfully on the 1.5m telescope at the Palomar Observatory in observing the astrometric binary Gliese 105A and detecting a faint red companion (Golimowski *et al.* 1995); also on the 2.5m DuPont telescope at Las Campanas observatory (see Figure 1.8). Below are listed some of the general science targets where an AO optimised coronagraphic capability would have an important application (Doel & Myers 1999):

- Brown dwarfs and sub-stellar companions (Nakajima *et al.* 1994).
- Protoplanetary disks around young stellar objects and associated jet phenomena.
- Circum-nuclear regions of AGNs and Seyfert galaxies.
- QSO host galaxies.
- Dust disks around main sequence stars.

- Symbiotic stars - studies of extended emission due to mass ejection events.
- Luminous Blue Variables (LBVs), study of their ejected material.

Instrumentation for additional scientific analysis down-beam from a coronagraph allows other new areas of research to be pursued. As listed in Table 1.1 there are currently only two coronagraphs that have this facility - CIAO on Subaru, which has a choice of linear polarimeters and OSCA on the WHT, which has a spectroscopic capability (with the OASIS integral field unit).

An important criterion in creating a high contrast imaging system is in keeping scattered light to a minimum. Uniquely, compared to other AO coronagraphs, the system at the WHT allows the insertion of an occulting mask before the AO system (see Figure 2.1 for location). Unlike the other masks in OSCA, this mask would ideally be made of a dichroic material which only allows the non-science (visible) wavelengths to pass through (these are used for wavefront sensing). If the observations were to be made at visible wavelengths then a neutral density or appropriate narrow band filter mask could be used instead to reduce the intensity of the central star, allowing through enough light for the wavefront sensor but still reducing the unwanted central starlight that contributes to scattering in the system. Unfortunately the many segments in the NAOMI DM contribute more scattered light than a similar sized continuous face-sheet mirror and the gaps between the segments also have a higher emissivity in the infra-red, so this additional pre-AO stop for the NAOMI-OSCA system is extremely useful in cutting down scattered light at the science wavelengths and thus improving sensitivity in the final image. Comparisons of sensitivity are made in §3.

As mentioned in §1.2, one of the ultimate aims for high contrast imaging systems in astronomy is to make a direct detection of an extra-solar planet. Comparing like with like, the best contrast systems will always be space-based observatories as they are not limited by atmospheric scattering. There are a number of space missions planned to detect planets, the two prevalent ones being Terrestrial Planet Finder (TPF) by NASA and Darwin by ESA. To give an idea of the challenges faced by a space mission to detect exo-planets, the required contrast ratio to detect an earth-like planet around a solar-like star at visible wavelengths, at a distance of 10pc is $\sim 10^9$ at only 0.1 arcsec from the star. In the mid-IR (8.0 - 30μm) this ratio drops to $\sim 10^6$, so most designs are based around

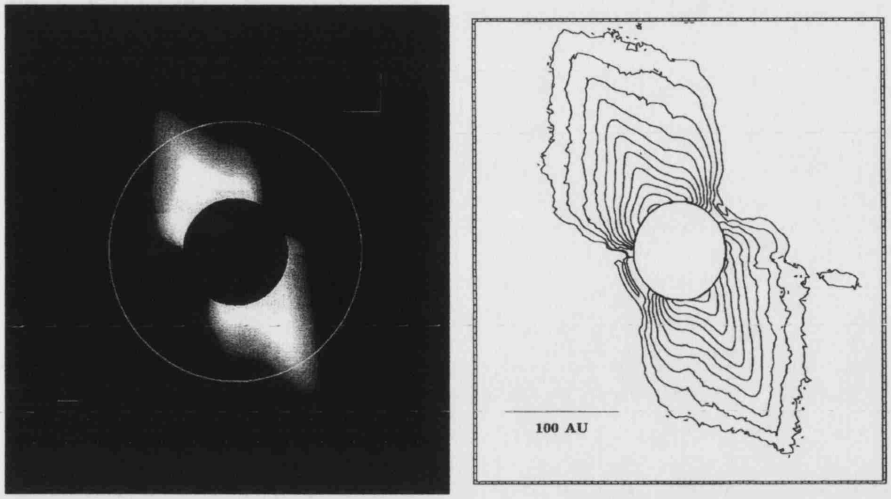

Figure 1.8: R-band image of the β Pic disk using the Adaptive Optics Coronagraph (AOC). Diameter of the occulted region is 5 arcsec and the superimposed circle is 12 arcsec, marking the change in disk brightness gradient and previous limit on disk photometry (left); Contour map of the R-band image in steps of 0.5 mag arcsec^2 from 18 to 13 mag arcsec^2 (right) (Golimowski *et al.* 1993).

these wavelengths. Ground-based designs are not constrained by size and weight as are the space designs, so plans for planet detection from the ground are centred around the idea of Extremely Large Telescopes (ELTs). Current designs range in size from 30 - 100m. In theory these would collect enough light to detect a planet at a reasonable integration-time but to obtain the necessary resolution depends on the development of what has been termed 'extreme AO'. These require an actuator density on the order of a few hundred per square meter of aperture area. The challenges in this field are enormous but by building on the knowledge gained from smaller systems we will be able to determine the best method by which these can be overcome.

CHAPTER 2

A Coronagraph for the William Herschel Telescope

A coronagraph facility for the William Herschel Telescope (WHT) was proposed by P. Doel and R. Myers (Doel & Myers 1999) mid-1999. The proposal was accepted and initial work commenced mid-2000. This coincided with the first commissioning of the Nasmyth Adaptive Optics system for Multi-purpose Instrumentation (NAOMI), which the coronagraph would be coupled with. This Optimised Stellar Coronagraph for Adaptive-optics (OSCA) had a rapid design development stage, with the preliminary design review in December 2000 and the critical design review in May 2001. The reviews are a necessity for a facility class instrument to ensure that the instrument conforms to the telescope's functionality (thermal properties, repeatability, lifetime issues etc.) and safety requirements. OSCA was finally commissioned at the telescope in May 2002 and became available for general use at the beginning of 2003.

The diagram in Figure 2.1 highlights the OSCA components (in blue), along the entire optical path at the GHRIL Nasmyth platform on the WHT. With reference to Figure 2.1, the path of light through the entire system is as follows:

Light enters the platform via the derotator optics which maintain the orientation of the stellar object as the telescope tracks across the sky. The beam comes to a focus in the alignment and calibration unit. This unit provides a number of different functions for AO set-up (white light pinholes - artificial star etc.) and is also the position where the

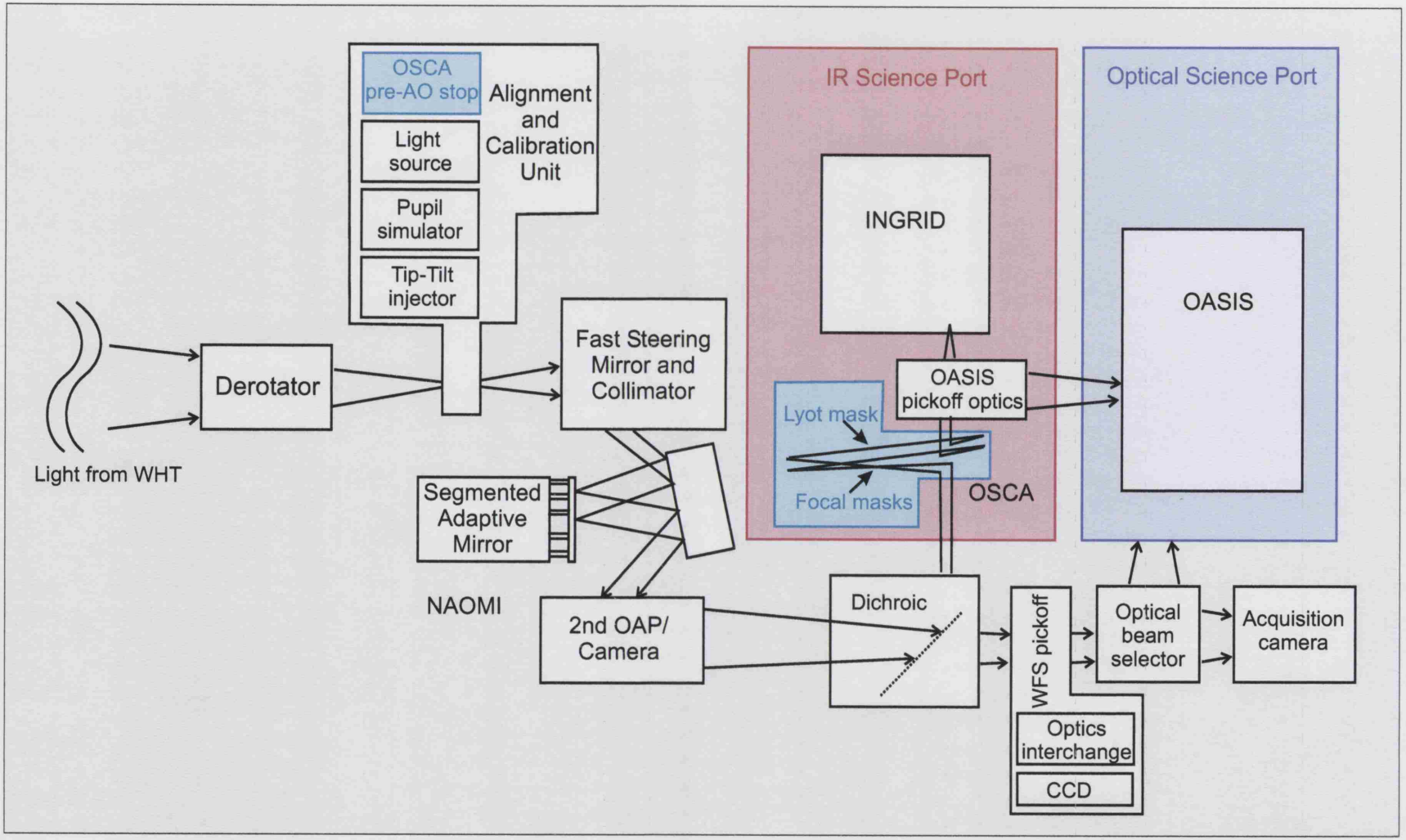

Figure 2.1: Optical sequence schematic of instrumentation on GHRIL table (modified from a NAOMI technical diagram (ATC 1997)). Sizes are not to scale and positioning does not necessarily indicate actual location of components on bench.

OSCA pre-AO occulting stop will be positioned. Next the beam encounters the first stage of AO correction, the tip-tilt mirror (called the Fast Steering Mirror) - this is a rapidly oscillating mirror (20Hz) which removes the average tip-tilt across the wavefront. The segmented mirror which is next along the path removes higher order aberrations. The beam is folded along the bench via the large pair of NAOMI OAPs. A large dichroic subsequent to OAP2 reflects the IR light ($> 0.9\mu$m) into the IR science port and transmits the visible portion of the beam to the NAOMI wavefront sensor. There are a selection of dichroics available depending on whether OASIS is used or not.

The transmitted visible light is passed to a Shack-Hartmann sensor comprising of a selectable set of lenslet arrays and CCD - information from this is fed into a computer control loop which determines the mirror positions that correct the incoming wavefront at both the tip-tilt mirror and segmented mirror. Also along this visible path is an acquisition camera and a small pick-off mirror which allows the guide star used for wavefront sensing to be steered on to the sensor. It is adjustments to the position of this pick-off mirror with the AO closed-loop that allows the star to be accurately placed behind the OSCA occulting masks (a short delay between adjustments gives the system the time lag needed for the AO position update to take effect).

The IR science port contains both OSCA (coronagraph) and INGRID (near-IR detector). With OSCA deployed the focus from NAOMI falls on the focal plane (occulting) masks in the OSCA system. The first OAP in OSCA parallizes the beam and in this parallel beam the Lyot (pupil plane) mask is positioned. A second OAP opposite the first brings the beam back to a focus within INGRID. With OSCA out of the beam the NAOMI focus passes unobstructed to INGRID. Pick-off optics between OSCA and INGRID allow the beam to be redirected to OASIS (an integral field unit), thus allowing the use of OASIS with OSCA for spectroscopy or as a visible light camera option.

2.1 Optical Design

An important requirement of the optical design was to ensure that the focal position was unaltered with OSCA deployed in the beam path and to fit this design into a very small spatial envelope (see Figure 2.2). The optical specifications were that it should be diffraction limited at a wavelength of 2.2μm over the full field of view and have less than

0.1 arcsec distortion over the field for the full wavelength range. The system also has a 1:1 magnification to leave the NAOMI field-scale unchanged. To accommodate the required wavelength range of 0.4 - 2.4μm all mirrors in the system were coated with protected silver on Zerodur. Similarly all transmissive optics were made from an infra-red (I.R.) grade fused silica (Spectrosil WF) which gives over 90% transmission across the entire specified wavelength range and can be polished to a very high surface quality.

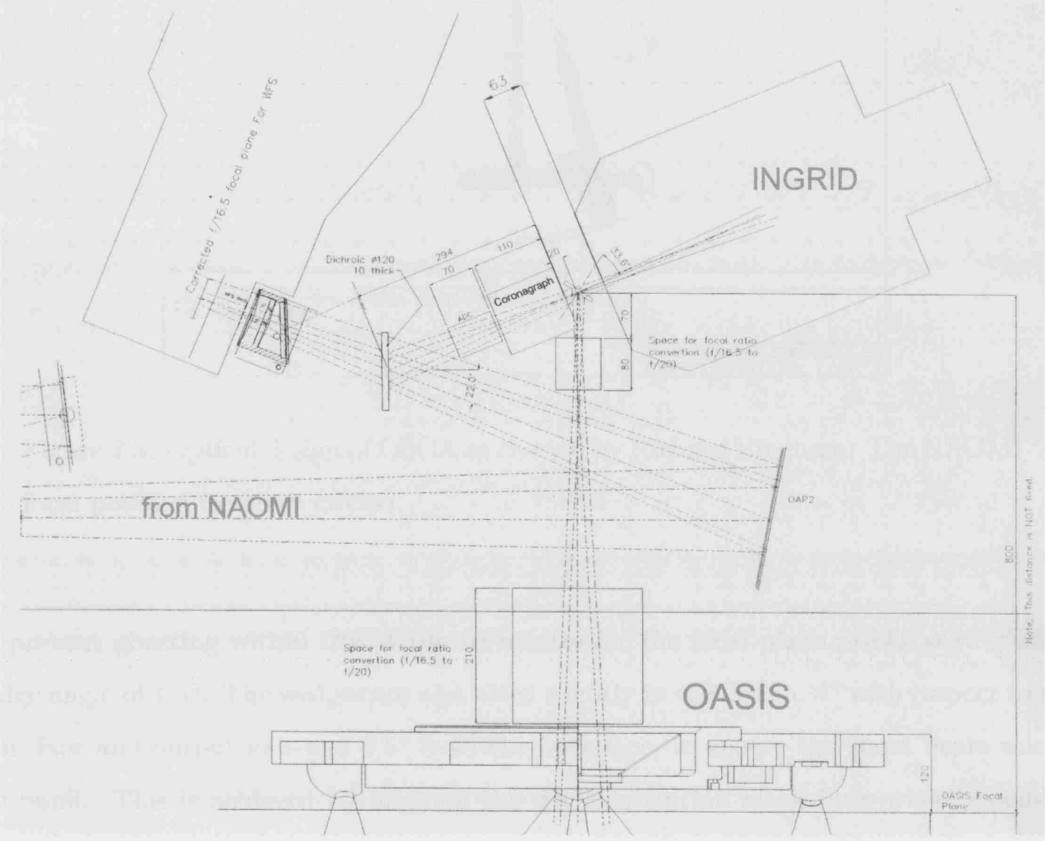

Figure 2.2: Original plans for the GHRIL optical bench layout showing the incoming beam from NAOMI and the position of future instrumentation, including that for the coronagraph (ING).

The field of view for OSCA is designed to be 20 arcsec. The image scale at the OSCA focus is 327.28μm/arcsec, so the substrates for the occulting masks had to be at least 6.6mm in diameter. To determine a suitable size for the substrates various other factors had to be taken into consideration: percentage clear aperture achievable on polishing (edge defects), edge overlap needed for mounting in filter wheel and finally, given the size of the filter wheel, the maximum substrate size that would fit 10 times round the circumference of

the wheel. Taking these factors into account a size of 12mm diameter was chosen for the substrates. Polishing gave a 90% central clear aperture (10.8mm diameter), and the clear aperture in the filter wheel is 10mm (to allow a 1mm ring all round for mounting).

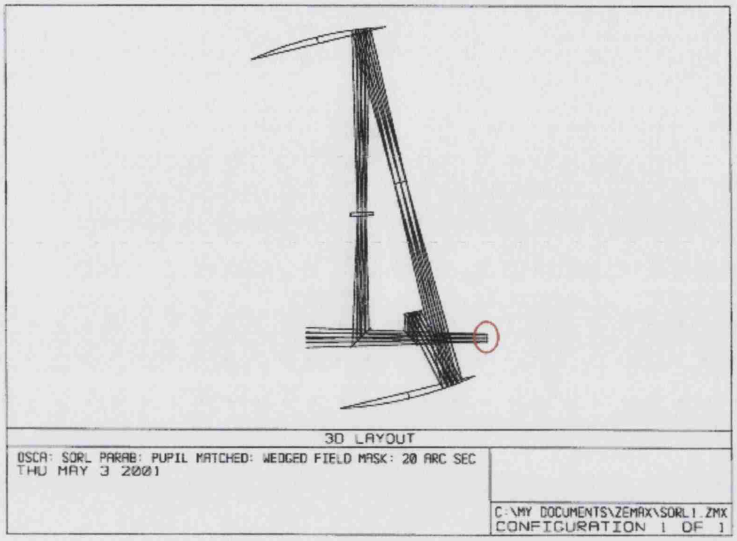

Figure 2.3: Optical design of OSCA as created by Richard Bingham. The NAOMI focal position is shown circled.

To prevent ghosting within OSCA the substrates for the focal plane masks were given a wedge angle of 1.5°. The wedges are also tilted slightly in the design, 4° with respect to the front-face and output axis and 2.5° from the back-face, to ensure the ghost beam misses the pupil. This is achieved by aligning the whole occulting mask filter-wheel assembly at a slight angle to the beam and checking that the ghost-beam is thrown completely out of the optical path. The edge of the substrates will contribute scattered light due to internal reflection and extra scattering will also result from surface chips and scratches gained during handling. To reduce this a razor-edged type anti-scatter mask (Figure 2.4) was designed that under-sized the field of view to 9mm and was installed just in front of the focal plane substrate.

Choice of material for the substrates was originally sapphire, since sapphire offers extreme surface hardness (not easily scratched), excellent I.R. transmission properties from 0.15 to 6μm and is chemically inert. Its higher refractive index of $\sim$1.76 also requires a smaller wedge angle than materials of lower indices. The optics company used to produce the substrates (Crystran Ltd.) informed us that sapphire is too hard to polish to the specifi-

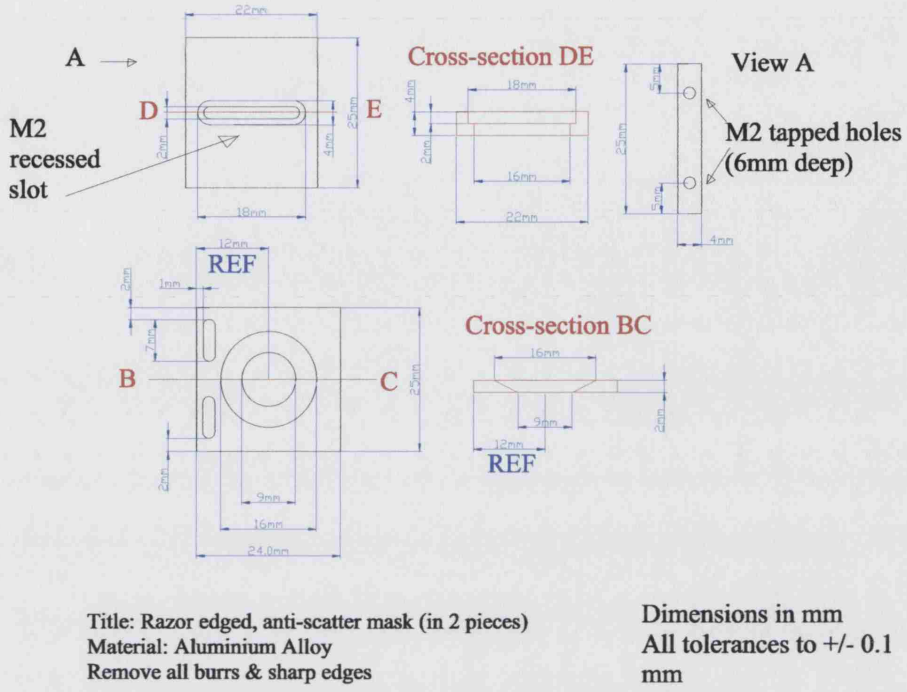

Figure 2.4: Design for the focal substrate anti-scatter mask

cations required and offered an alternative material - Spectrosil WF. This is a water-free, synthetic grade fused silica which has a transmittance of greater than 95% up to a wavelength of 2.4μm as shown in Figure 2.5. The refractive index is 1.47 - 1.43 for the 0.4 - 2.4μm wavelength range. The main compromise is in surface hardness which is about an order of magnitude lower than that for sapphire and so the Spectrosil substrates are more easily scratched. The final specifications for the substrates are as follows:

- Over 90% transmission for the wavelength range 400 - 2500nm
- Diameter = 12mm $\pm$ 0.1mm (at least 10mm clear aperture)
- Central thickness = 3mm $\pm$ 0.1mm
- Wedged angle = 1.5° $\pm$ 0.05°
- Surface flatness = $\lambda/20$
- Surface quality = 40:20 scratch/dig

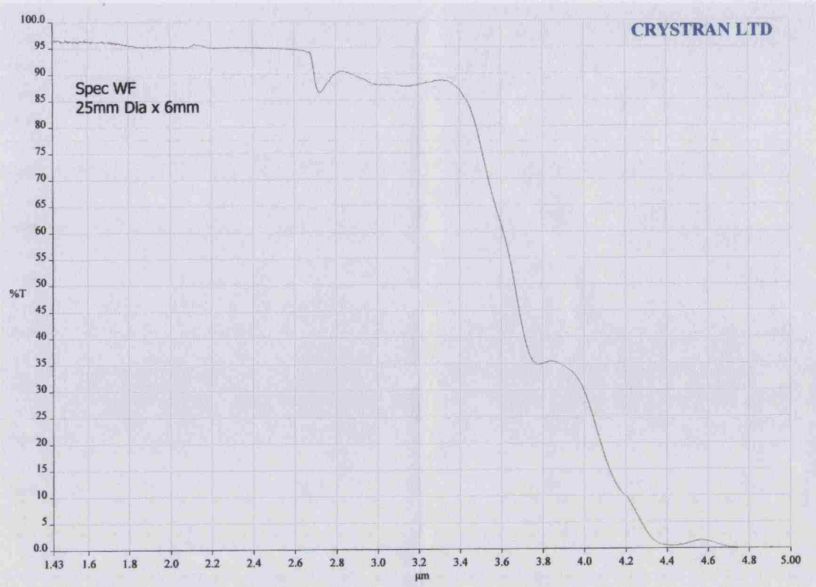

Figure 2.5: Percentage transmission versus wavelength for a 6mm thick sample of substrate material.

- Both sides uncoated
- Apex of angle marked on each window

Since these substrates were made to order, delivery was approximately 10 weeks and 12 pieces were ordered initially to allow for any possible damages in the photolithography process.

The choice of material for the mirror substrates was Zerodur – a ceramic with a near-zero coefficient of thermal expansion and the usual choice for a $\lambda/20$ surface requirement. The reflective coating was chosen to be protected silver (also known as over-coated silver), with a reflectance, $R_{avg} > 95\%$ over the specified wavelength range. A protected coating was chosen to prevent oxidation (components need to be long lived) and any surface oxidation would lead to reduced reflectance and increased scattering over the entire spectrum which is very detrimental to a high-contrast imaging system. Protected silver offers the most durable and highest reflectance coating over the required wavelength range compared to other coatings such as aluminium. There are a total of five mirrors in the OSCA system – 3 flats and 2 off-axis paraboloids (OAPs). The flat-mirrors can be bought 'off the shelf' to almost any specification: the flats in OSCA were purchased from Melles Griot and were 25mm in diameter, 6mm thick with $\lambda/20$ surfaces. The OAPs however were very specific

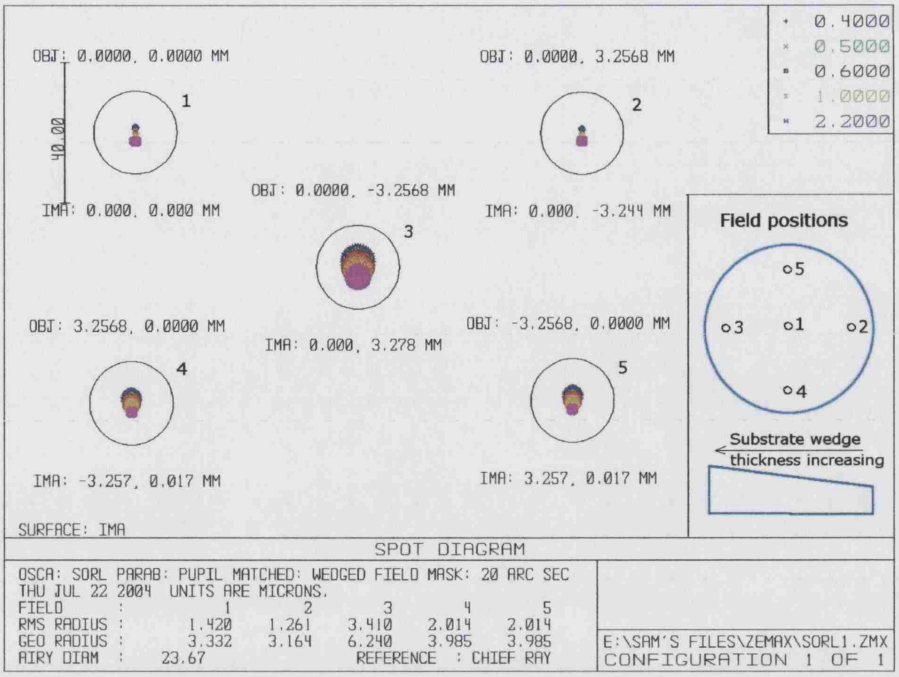

Figure 2.6: OSCA spot diagrams at five field positions (centre, up, down, left and right), shifted by 3.2568mm ($\sim$ 10arcsec) from the centre in each case. The black circular outline about each is the size of the Airy disk (1.22λ/D) at a wavelength of 0.6μm.

and a company had to be found to produce them in a limited timescale.

Specifications for the off-axis paraboloids:

- Diameter = 30 - 35mm
- Edge thickness = 16mm $\pm$ 1mm
- Clear aperture = 90% diameter
- Off-axis angle = 14 degrees
- Focus to surface = 157mm $\pm$ 0.75mm
- Figuring tolerance = 0.1 waves P-V (visible)
- Surface quality = 60:40 scratch/dig
- Two cut from one master

- Master substrate diameter = 117mm ± 1mm

The Space Optics Research Laboratories (SORL) were chosen to produce the OAPs, however due to severe shortages of optical polishers the estimated production timescale of 8-12 weeks turned into 6 months and was the biggest delay in commissioning for the instrument.

The final optical layout for OSCA is shown in Figure 2.3 and the spot diagrams for five different field positions in Figure 2.6. This is the third revision to the design; initially the design only had 2 flat mirrors, however changes in space limitations on the instrument bench led to a third flat mirror having to be included. Original specifications (not given here) for the off-axis paraboloids also had to be changed slightly to shorten manufacturing/delivery time, resulting in further alterations to the design.

2.2 Coronagraphic Masks

OSCA is based on the classic Lyot (Lyot 1939) coronagraph design. The sizes of the focal plane masks were chosen to take advantage of the improved PSF offered by NAOMI. Ten focal plane masks are available for selection on the instrument. One is for alignment purposes only and two are of a more novel gaussian profile compared to the standard hard edged circular disks.

2.2.1 Standard focal plane masks

The standard focal plane (occulting) masks are made of chromium, these were deposited on the substrates by a contact-photolithography process. To manufacture the master template that is used in this process, a CAD drawing of the required masks was sent to the Central Microstructure Facility at the Rutherford Appleton Laboratory (RAL).

The dimensions of the masks that were included in the CAD file are listed in Table 2.1. The sizes were chosen to be in units of $1.22\lambda/D$ at the longest and optimum wavelength for the adaptive optics system (2.2μm), with the smallest mask sized on the width across the first minima of the Airy disk. The target alignment stop, also in the OSCA focal-mask wheel, is shaped like a target sight and is detailed in Figure 2.7. The width of the ring and cross-hairs were picked to match the pixel size in INGRID. This target provides a good

Table 2.1: Size range of occulting masks on master template, denoted in units of μm and resolution at minimum and maximum wavelengths

Mask #	Ø in λ/D @ 2.2μm	Ø in λ/D @ 0.4μm	Ø (μm)	Ø (arcsec)
1	2	11.10	86.69	0.265
2	3	16.60	130.04	0.397
3	4	22.10	173.38	0.530
4	5	27.75	216.73	0.662
5	6	33.16	260.07	0.795
6	8	44.21	346.76	1.060
7	12	66.31	520.14	1.590
8	15	82.89	650.18	1.990
9	Target alignment stop		327.28	1.000

focus aid when aligning the mask-wheel assembly and can be used as general alignment tool for viewing on INGRID when otherwise the occulting masks would block the light beam.

To save on costs and number of runs throughout the photolithography process, all the solid masks were created on a single 100mm x 100mm plate template. This plate along with a specially made holder can be seen in Figure 2.8. The production process at RAL involved taking a plate of the chosen size that had a near perfect thick coating of chromium. The

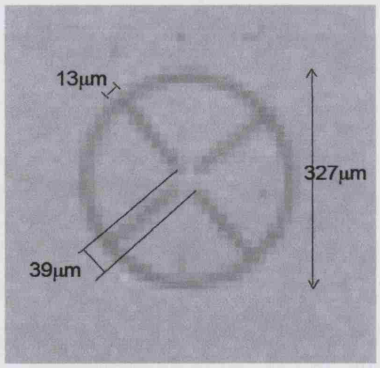

Figure 2.7: Picture of target mask taken with INGRID. 1 pixel $\simeq$ 13μm on INGRID

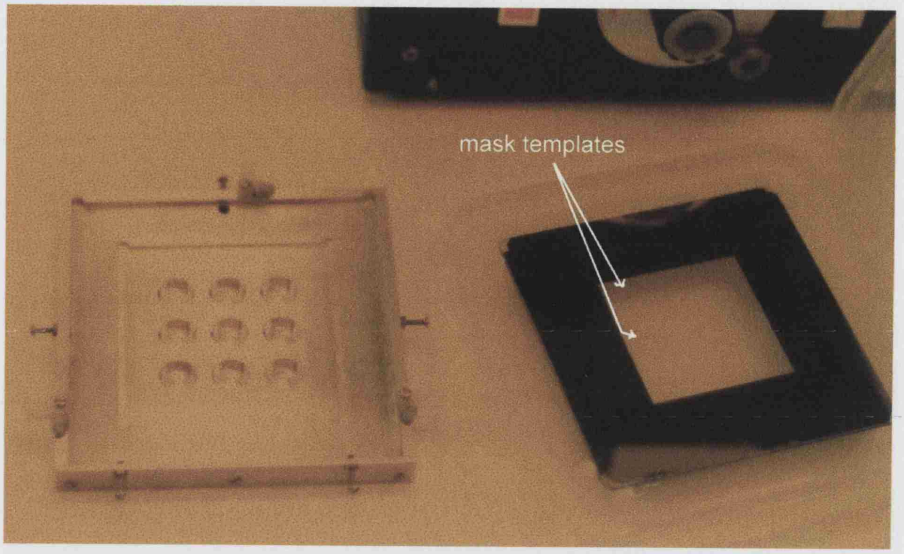

Figure 2.8: Photograph of the plate holder and master template plate made by RAL, opened in the UCL clean room. The two largest mask designs can just be seen on the template.

plate design is loaded into a computer controlled electron-beam which 'writes' the design on to the plate - the chromium coating is vaporised where the e-beam hits. For a mostly light-field design such as for the OSCA masks this can be a lengthy procedure. The remaining chromium that was not part of the design can be seen around the edges of the plate.

To enable easy and accurate use of the master plate template a plate holder was also designed, Figure 2.9. This housed the 12mm substrates in the appropriate configuration and allowed the mask template to be positioned with an accuracy of approximately ± 1 mm over the substrates so that the occulting spots would be over the centre of them.

To allow for the wedge of the glass substrates and to present a flat surface for placing the template upon, a thin disk of foam was placed at the bottom of the wells in the plate holder and then the glass substrates laid on top.

Photolithography process for occulting mask production

To ensure the photolithography process and the plate holder set-up was suitable for producing acceptable, centred occulting spots on the substrates a test run was carried out

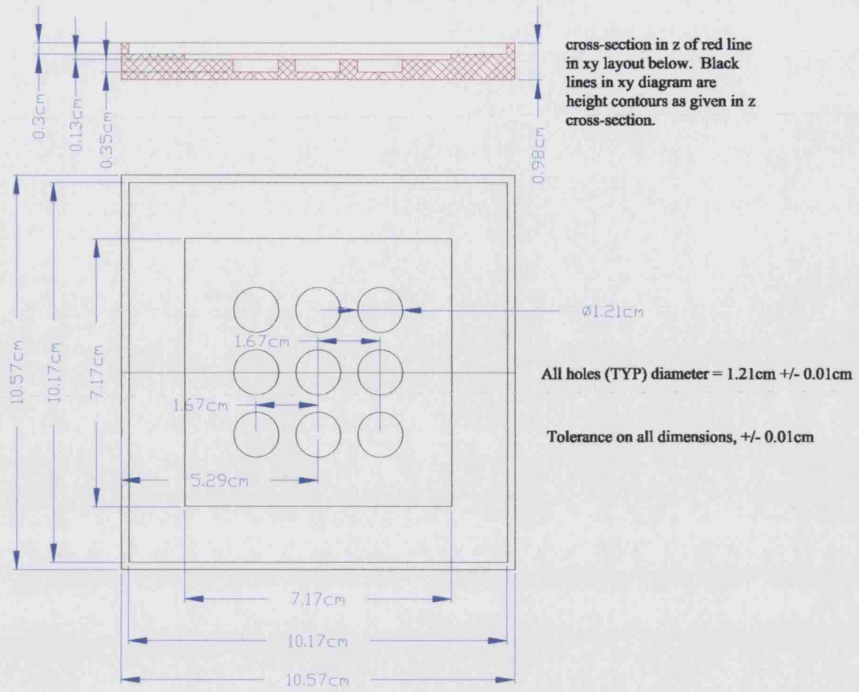

Figure 2.9: CAD drawing of the holder designed to hold the mask substrates and RAL photolithography plate. This holds them in the correct position during the UV exposure step of the photolithography process, see Figure 2.10(d).

using some cheaper test substrates. These were created by gluing together three 10mm × 1mm flat round windows and using a piece of paper during gluing to create a slight wedge of ~ 0.1mm height difference either side along the apex diameter which approximates that of the real substrates. A 100nm layer of aluminium was then deposited on the test masks. Each substrate was placed in turn on a rotating wheel and several drops of photoresist were used to cover the aluminium layer, the excess was spun off to leave a very thin layer of photoresist. The substrates were then placed on a hot-plate (100°C) for a few minutes to bake on the photo-resist layer. The mask-holder was lined with small circles of thin foam (to take up the wedge and present a flat top surface) and the substrates carefully placed into the holder so that they were well centred. The RAL master template was then lowered on to the masks and held down with 4 small plastic clips round the edge of the plate holder. Interference fringes are observed when there is good contact between the substrates and the template. A UV light box was then placed over the assembly and exposed for a short time. The UV light degrades the photoresist polymer that is not covered

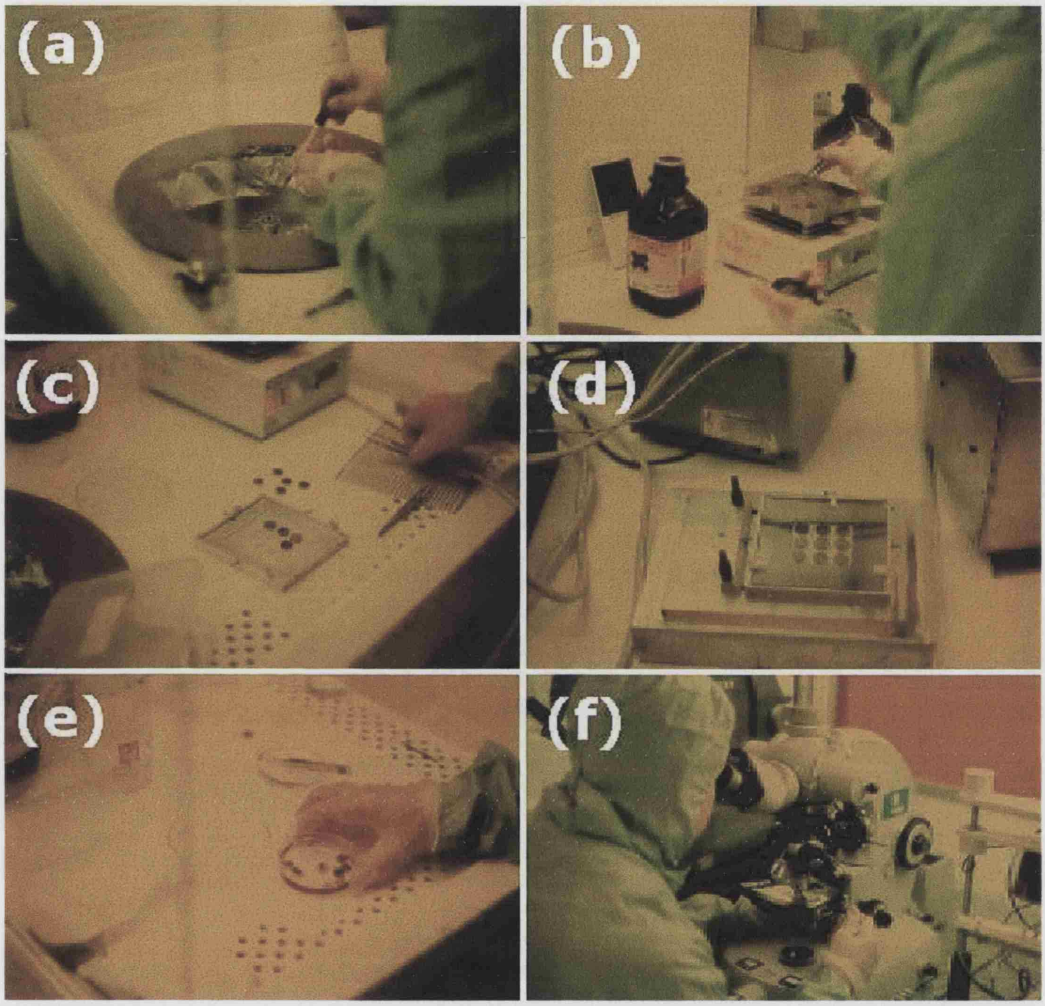

Figure 2.10: Photographs of the various stages in the photolithography process at UCL used to create the occulting masks. (a) spinning the photo-resist onto the substrate; (b) baking on the photo-resist layer; (c) placing the substrates into the master plate holder; (d) substrates + RAL plate on top for UV exposure; (e) chemically dissolving the exposed photo-resist and chromium layer underneath; (f) checking the final masks.

by the mask pattern. The substrates were then individually placed in developer solution (~14 seconds) to develop the photoresist layer (like photographic film). The substrates were then observed under a microscope to check the integrity of the pattern. At this stage during the first lithography run it was discovered that the developing solution caused the aluminium layer to bubble up and thus it was easily rubbed off. This is one of the reasons chromium was chosen over aluminium for the final masks. The aluminium layer that had been exposed to the UV light was then etched off using phosphoric acid to leave just the mask pattern in aluminium on the substrates. The timing of this stage is crucial since if left too long in the acid it will start to undercut the photoresist layer and dissolve the mask pattern. The masks were then checked again under the microscope and then the remaining photo-resist layer was removed using acetone. The stages in this process are summarised in Figure 2.10.

For the production run we initially used aluminium to make the occulting masks but we encountered some difficultly using this to make such small masks. These included the layer of aluminium on the substrate bubbling during the photolithography process, frequently leading to small holes and tears in the mask, poorly defined edges and poor adhesion of the aluminium to the substrate so that even during a very light cleaning process all or part of the mask would be rubbed off. The reason for initially wanting to use aluminium was to allow the focal plane masks to have a small, non-zero transmission (about 0.1%). This would allow easier centring of the mask over the star during observing runs. Experiments were carried out to find out the thickness of aluminium required for such a transmission. The evaporation process however has an accuracy of $\pm$ 5nm in the thickness of aluminium that is deposited. The reason for this is that the layer thickness is sensed by an oscillator (~6MHz) that changes impedance with the changing thickness of aluminium on it. However since a 50nm layer only takes about 5 minutes to deposit, laying down thin layers becomes very inaccurate. An experiment was carried out using glass microscope slides with different thicknesses of aluminium deposited on them. The transmission of these plates at red laser wavelength (613nm) was tested using a medium area photodiode with a large response range (input range from $\sim$ 0.0001 mW/cm^2 to 100 mW/cm^2) and an oscilloscope. Results indicated that changes in thickness smaller than the depositing error around the thickness of interest, would lead to large changes in the transmission at visible wavelengths. It was decided that this method was too inaccurate

to produce the thickness that would give the required transmission and so we decided to make the masks from a thick (~100nm), opaque layer of chromium.

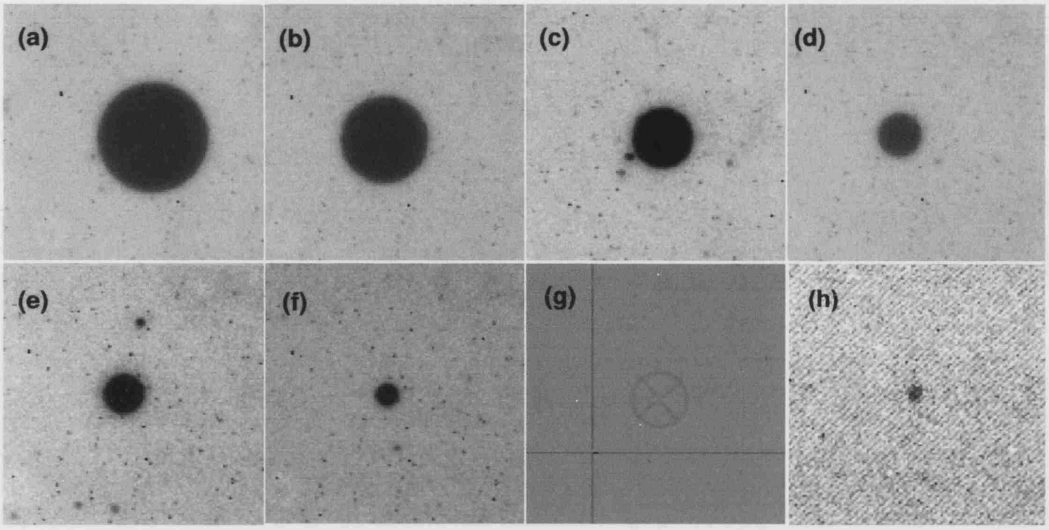

Figure 2.11: The hard-edged, chromium deposited focal-plane masks in OSCA. (a) 2.0 arcsec; (b) 1.6 arcsec; (c) 1.0 arcsec; (d) 0.8 arcsec; (e) 0.6 arcsec; (f) 0.4 arcsec; (g) 1.0 arcsec target alignment stop; (h) 0.25 arcsec stop - the poor image quality is due to read-out problems on INGRID, characterised by the diagonal striping. Illumination is provided by the standard room lights in GHRIL. All sizes are on-sky diameters.

Chromium is also less likely to bubble due to better bonding/adhesion to the glass surface on deposition and thus it is also more resilient to cleaning and scratches. Another benefit is that chromium etches at a slower rate than aluminium and so is less prone to undercutting since there is a longer period of time in which to remove the substrate from the acid. A 130nm thick layer of chromium was deposited on the substrate for the final production run. Three runs were needed to produce 8 acceptable masks (Figure 2.11). This is because with photolithography there is a diminishing return each time a plate is reprocessed; the surface of the glass is spoiled by tiny deposits of chromium that cannot be removed and damage by scratches and pits obtained from the handling and chemical processes. On the last run, glass plates that had been through the process had visible scratches when viewed under a microscope at $\times 8$ magnification. This is the main reason why the 0.5 arcsec occulting mask as listed in Table 2.1 was not produced: by the last production run a satisfactory mask of this size had not been created.

2.2.2 Gaussian focal plane masks

Computer simulations, both here (in §3) and by others (Nakajima 1994), have shown that occulting masks of a gaussian profile give better suppression than the standard solid disk design. However creating such a mask is not trivial. A number of methods of producing a gaussian-like mask were investigated and in-house experiments were performed over several months but with limited success.

One idea was to produce a photographic glass plate using the established technique of wet-plate photography. This involves pouring a solution of collodion – a mixture of raw cotton, nitric and sulphuric acid, ether, alcohol, iodide and bromide – over the glass substrate, which sticks to the surface to coat it completely. This is then dipped into a silver nitrate solution which reacts with the coating to form a photo-sensitive silver halide coating. The substrate could then be exposed to a gaussian-profiled beam of the required width to create the occulting mask. A developing solution of iron sulphate and acetic acid is then poured over the exposed plate which turns the exposed silver-halide grains into metallic silver. After washing, the unexposed silver-halide is removed by washing in a tray of sodium thiosulphate which also acts as a fixing agent. The substrate is then washed and dried and traditionally coated with a varnish. Although this is would be unsuitable for good optics, it was envisioned that varnish could be replaced by a suitable anti-reflection coating instead. The coating is required to protect the otherwise delicate silver image. Due to unsuitable facilities and short remaining time-scales in production this method was unable to be tested. The lifetime, robustness and optical density of masks created this way is unknown and so would have had to undergo thorough testing to determine these factors. However if successful it would mean that a gaussian-type mask could be created that would be suitable for any wavelength application since silver has a very broad-band reflectance and can be deposited on any suitable transmissive material.

Another idea that was tested experimentally was to create a graded-edged mask by evaporating metal (aluminium or chromium) through a small hole on to the substrate. The experimental set-up is illustrated in Figure 2.12.

Brass shims with thicknesses ranging from 0.2 - 0.9mm were cut to match the size of glass slides (75 × 25mm). Two holes of diameter 1mm were drilled into each of the shims. Small strips of brass shim were then araldited to the underside of these plates to give

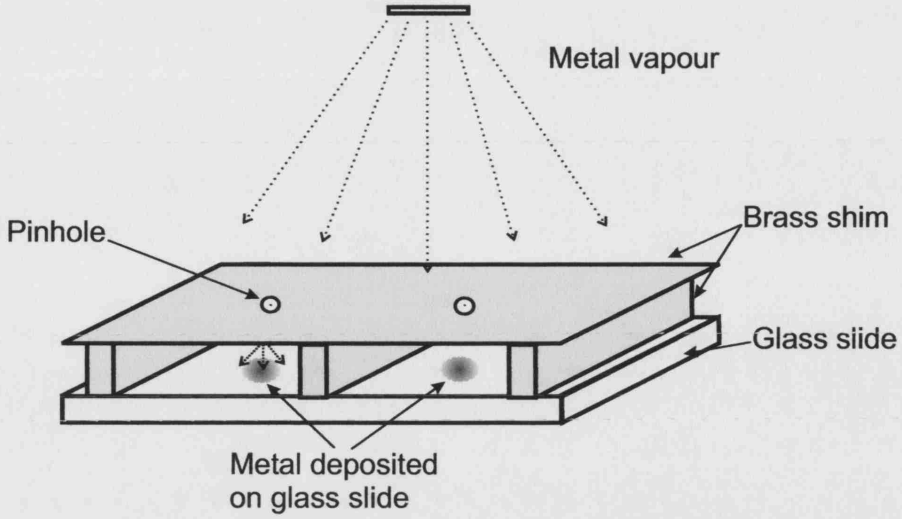

Figure 2.12: Experimental set-up to produce graded-edged occulting masks. Sizes are exaggerated to illustrate process.

the required spacing between the holes and the glass plate. Unmounted pinholes (3mm diameter, aluminium rounds) were then carefully araldited over the drilled holes in the shim so that the pinhole was roughly centred. Once the glue had cured, the set-up was carefully cleaned (including the glass slides) to be free from grease and dust and the two parts – brass shims + pinholes and the glass slide – were sealed together using masking tape on all sides and around the middle to ensure even contact. The combined unit was then placed in the metal evaporation chamber and given several coatings. A clear glass plate was also placed in the chamber for an indication of the layer thickness. The experiment was repeated for a selection of pinhole diameters (12.5, 25, 50, 100 and 200μm), gap heights between the pinhole and the glass slides and a choice between aluminium or chromium for the metal vapour. Metal deposits on the glass slide were only recorded when using the 100μm (only just visible by eye) and 200μm (visible, but extremely thin deposit) pinholes, and a gap height of $\sim$0.4mm. It also appeared that chromium gave a higher throughput through the pinholes than aluminium. There could be several reasons for this difference such as the vaporising chamber set-up (e.g. level of heating current) or other properties which were non-quantifiable with the experiments performed. Creating a high enough throughput to the glass plate to achieve the required density of metal for an occulting mask was the biggest problem in this experiment. Due to time restrictions this method was abandoned, although further experiments might yield the desired result.

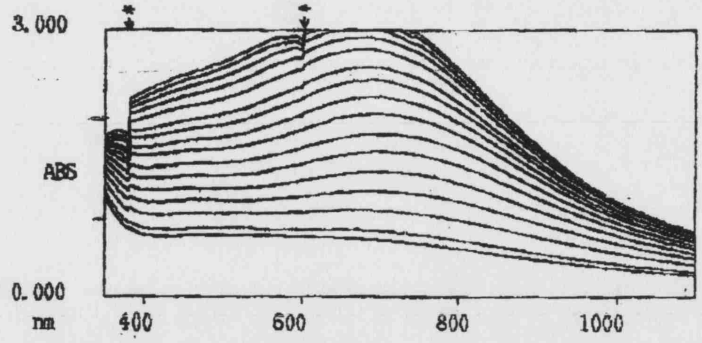

Figure 2.13: Transmission spectra (ND values) of HEBS-glass GI85 plate at various e-beam exposures. Glitches in the graph occur where a change of grating was used and is marked by a $\star$. Data supplied by Canyon Materials Ltd.

An extensive internet search for companies that could produce a gaussian type mask turned up only one with a suitable product. Canyon Materials Ltd have patented a glass formula, HEBS (High Energy Beam Sensitive)-glass, which behaves in a similar way to photographic film. HEBS-glass is sensitive towards e-beam exposure where exposure with a certain electron beam dosage changes the optical density of the material. In this way the HEBS-glass can act as a mask material and by varying the e-beam exposure different greyscale (optical density) levels can be written into the glass. To create the gaussian profile masks for OSCA (Figure 2.14) 669 different greyscale levels were written in ring steps of 0.5μm width. The smallest greyscale resolution element that can be written with this process is 0.1μm across, however this is extremely costly. A table was provided to Canyon Materials listing the percentage transmission for each of the 669 rings that make up the mask. These were obtained by picking points along a gaussian curve with a full-width half-maximum of 0.5 arcsec (165μm). The possible optical density values that can be written are in the range 0.07 - 3.0, calibrated at a wavelength of 760nm. Canyon Materials provided (free of charge) an alternative mask profile to the one specified by adding 0.07 to all the grey levels, and where this exceeded 3.0 the value of 3.0 was used, i.e. the maximum optical density was used for the inner 4μm radius of the mask. This produces a smoother edge profile for the mask - this is labelled as the Gaussian mask B. One restriction on their usage is that the glass is only suitable for the wavelength range 0.4 - 0.8μm, Figure 2.13 shows the variation in optical density over these wavelengths. This is the company's longer wavelength formula, their standard formula (for most laser

applications) is designed for 0.3 - 0.7μm wavelengths. The maximum optical density at 800nm is 2.5 (compared to only 1.0 for the standard formula), and continues to drop with increasing wavelength so the masks cannot be used for near-I.R. observations (as with INGRID). There is a visible camera on OASIS, so these masks were included in OSCA with this use in mind. The HEBS-glass photomasks are non-reflective (~4%), and have near zero difference in reflectivity between darkened and undarkened areas. Since the mask is created within the glass surface rather than on top (as in a photolithography process) the masks are much more robust to damage, cleaning processes etc. and cannot be rubbed off.

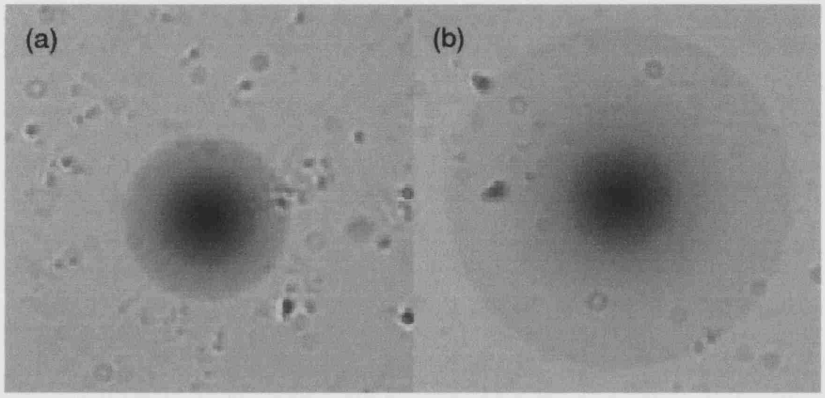

Figure 2.14: HEBS greyscale gaussian masks imaged with a SBIG camera. The small blotches are due to dust on the CCD window. (a) Gaussian A - measured FWHM of 0.50 arcsec. (b) Gaussian B - measured FWHM of 0.56 arcsec. The contrast level in each image is different to best bring out the different greyscale levels.

Since these masks are written on flat, 5-inch glass plates they had to be core drilled out to 12mm diameter circular pieces and then bonded to blank Spectrosil windows to give the required wedge. The two pieces were bonded together with an optical quality glue that best matched the average of the refractive indices. This was Norland 81 (NOA 81) UV curing glue. When cured, NOA 81 has a refractive index of 1.56 and a spectral transmission curve of ~95% across the wavelength range 0.4 - 2.0μm, dropping to 90% at 2.5μm. The extra 2mm glass thickness produces a negligible beam shift.

2.2.3 Pupil plane masks

The pupil plane masks are manufactured from 0.25mm thick, hard stainless steel by a process of photo-chemical machining. Unlike the occulting mask assembly, where a choice of masks can be automatically dialled up from the control room, the Lyot stop assembly was designed only for manual change of the mask. The main reasons behind this were a lack of space on the instrument bench, cost savings, ease of use and the fact that a single size of Lyot stop can be designed that is optimal for most observing applications. The optimal size is essentially a trade-off between system throughput and suppression. The final image quality must also be taken into account, since the Lyot stop is acting as a spatial filter. A Lyot stop that cuts out too much of the light in the pupil plane will heavily modify the PSF shape and fine structure in the final image will be lost.

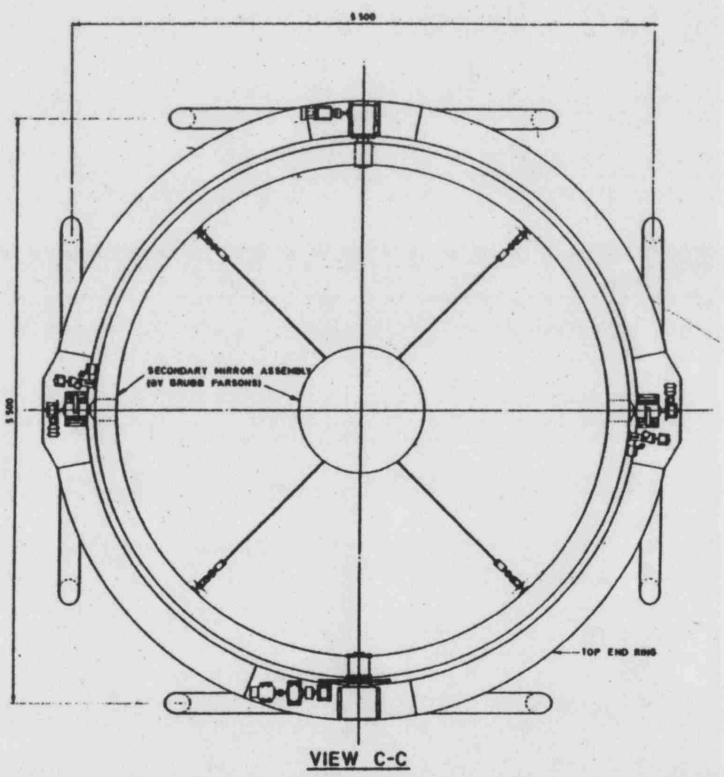

Figure 2.15: Diagram of the view down the WHT mirror-tube assembly (RGO/ING).

Exact measurements for the shape of the Lyot stop were taken from the original CAD diagrams for the WHT mirror tube assembly, see Figure 2.15, and confirmed by the ING.

The measurements were then converted into the sizes at the OSCA pupil plane, taking the size of the pupil from the Zemax optical design to be 9.8mm. Computer simulations (as

Table 2.2: Dimensions of the telescope aperture components in the OSCA pupil-plane

Component	Actual size (mm)	Size at OSCA pupil (mm)
Primary mirror	4200	9.80
Secondary mirror cell	1174	2.74
Telescope vane width	25	0.058
Vanes, offset from centre	100	0.23

described in §3.1.1) helped determine the Lyot stop size which was chosen to be 80% of the primary and 120% of the secondary, as in Figure 2.16. Masking for the telescope vanes was also included in the Lyot stop design and these were chosen to be 6 times oversized to the vanes. An additional Lyot stop was manufactured of size 90% primary, 110% secondary which is an optimum size for shorter (visible) wavelengths or in the case where a higher throughput is necessary. The percentage throughput compared to the full aperture for the (80,20) Lyot stop is 53% and for the (90,10) stop is 72%, these numbers take into account the over-sized vane masking.

An important consideration in the size of the vane masking was the smallest size that the chemical etching/machining process is capable of. This minimum size is approximately 10% of the thickness of the metal used to make the mask. The thickness of the Lyot stops was chosen to be 0.25mm to ensure a rigid mask: a thickness below this could have been too flimsy. Based on this the minimum size/precision of the machining process was 0.025mm. Although a photolithography process on glass would allow almost any size, the number of optical surfaces in the system is best kept to a minimum and so the photo-chemical etching process was used instead. Ten masks of each size were made since the cost was the same as for one. This allowed for plenty of spares if masks are misplaced during manual changes.

Additional pupil stops were manufactured at the ING to go in the INGRID pupil filter wheel. This provides cooled masks for OSCA which are particularly important for observing in K-band otherwise the thermal background noise is very high. Two stops were made

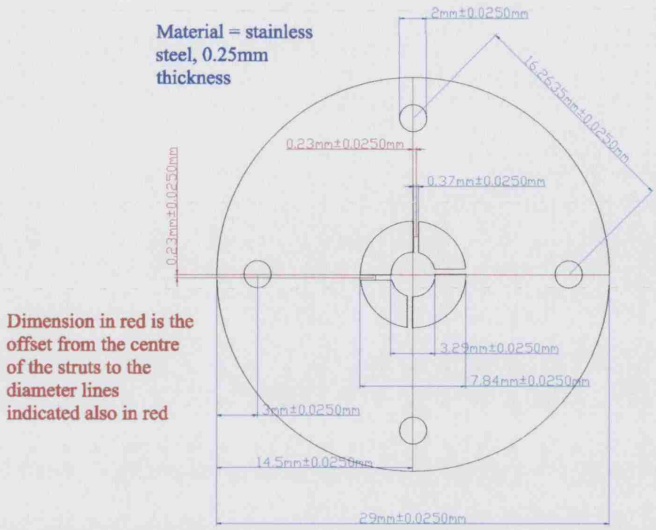

Figure 2.16: Lyot mask design for the (80,20) size, as sent to Microponents for manufacturing.

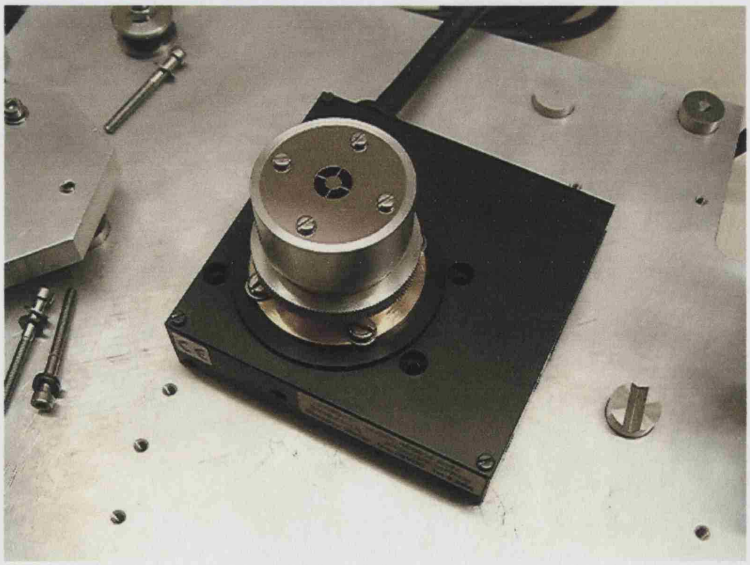

Figure 2.17: Photograph of the Lyot stop mounted on the SR50 rotation stage.

named OSCA_OBS and OSCA_CLEAR. OSCA_OBS consists of 80% full pupil masking with a central obscuration 115% of the normal obscuration size supported by three vanes which are as thin as possible. Since the INGRID pupil size is 4mm, this corresponds to a mask of outer diameter 3.2mm with a 1.387mm diameter central obscuration supported on vanes 150μm thick. OSCA_CLEAR is just an 80% pupil mask with no central obscuration. Unlike the OSCA pupil masks which rotate with the telescope vanes, the INGRID masks are fixed in position.

2.3 Mechanical Design

The important mechanical requirements of the OSCA system were for it to be thermally stable from -5°C to 25°C (i.e. positional accuracy is held within the specified tolerances and that the components will function correctly), vibrationally stable, for none of the components to infringe upon the beam path within or outside of OSCA, the entire system to be deployable in and out of the beam with a repeatability of $\pm$50μm, the centre of the focal plane masks to be automatically positioned to $\pm$10μm, 5 year component lifetime (standard ING requirement) and for the heat output to be minimised so as not to interfere with INGRID.

The design of OSCA meets its performance requirements by using kinematic referencing wherever possible to provide precise adjustment and repeatable positioning of the active parts of the system. Commercially available components were used where possible.

Structurally, OSCA consists of two horizontal plates with the opto-mechanics sitting on the top plate. The lower horizontal plate sits on the optical table. It is positioned by screw adjusters and clamped rigidly to the table. The top plate is spaced from the lower plate by three vertical fine thread ball-tipped screws with kinematic location, which provide height and tilt adjustment for the complete optical system.

The unit can be automatically deployed to interrupt the beam passing through to INGRID and the OASIS pick-off mirror. Space restrictions on the GHRIL table only allowed for a vertical deployment system and using the existing kinematic system provided a simple and reliable solution. Removing OSCA from the light path is achieved by lowering the top plate at the incoming beam end. The flat and ball-tip adjuster separate, allowing the top

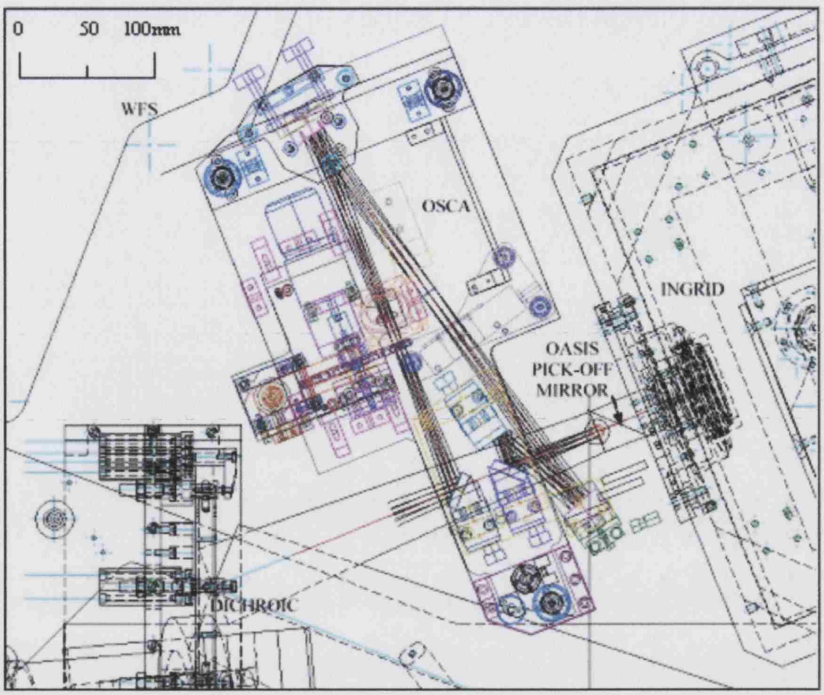

Figure 2.18: CAD drawing of OSCA with reference to surrounding instrumentation (Andy Charalambous)

plate to rotate about the groove and cone. The top plate by itself is loose. It is held in place by a linear pneumatic cylinder (piston) which links the top plate to the end bridge that is bolted on to the bottom plate. The pneumatic cylinder acts as the actuator for deployment.

OSCA is out of the beam in its inactive state. To deploy into the beam the pneumatic cylinder is air pressurised which compresses a spring mechanism as it lifts the top plate up and is stopped against a pad of hardened steel. Two reed switches attached to the pneumatic cylinder at either limit of the piston motion provides the deployment status signals.

The occulting mask wheel is purpose-built for OSCA as dictated by space and performance requirements. The wheel holds 10 masks that are glued into position; the detailed wheel design is shown in Figure B.7. The mask wheel is rotated by an on-axis stepper motor and encoder, the encoder being used to confirm mask position. Figure 2.20 illustrates how the pneumatically actuated pivoting detent arm locates in V-notches on the wheel circumference to provide a high level of accuracy and repeatability. Since the forces of

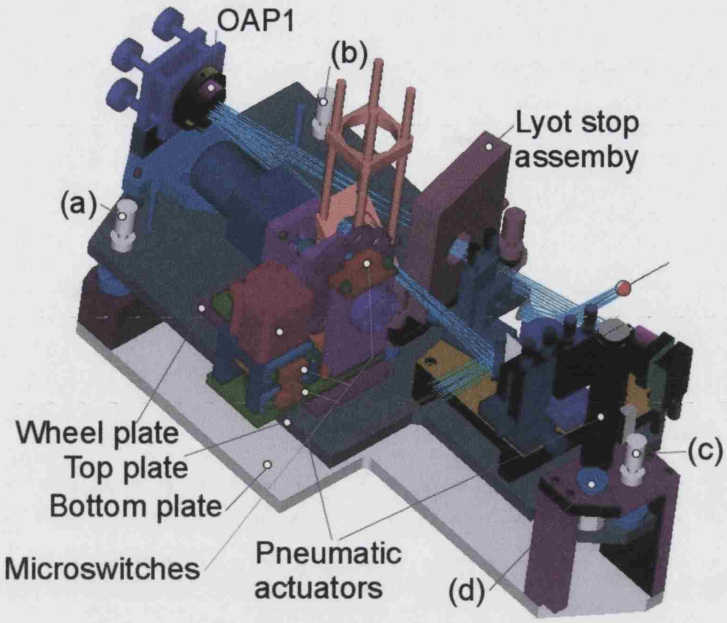

Figure 2.19: 3-d CAD construction of the OSCA system (Andy Charalambous). Adjustable, ball-bearing tipped locking screws - (a), (b) and (c) provide precise height positioning of top plate and additionally (a) and (b) rest on groove and cone pads (light blue disk under (a)) to provide the pivot point for the deployment mechanism. Screw (d) provides a manual mechanism to lock the top plate up if needed (e.g. air supply failure).

the detent arm act through the centre support of the wheel, the wheel should not move out of alignment with the repeated action. Once the detent arm has locked the wheel in position the stepper motor is turned off; the beam passes very close to the motor so having the motor turned off during observing ensures minimum contamination of the beam from heating effects, and removes any negative effects of heating on the AO system. The co-alignment accuracy of masks is achieved by using the actual wheel unit and detent arm during assembly. Once the first mask is glued in position an optical reference of its centre is made and the remaining masks are glued so that they are centred on this reference. This eliminates any errors due to machining inaccuracies. A metal rod attached to the detent arm activates a microswitch above or below it when the arm moves up or down, this provides the signal to indicate whether the detent is locked in or out. The wheel rotation mechanisms are disabled when this signal indicates the detent is locked in. The wheel's home position is provided by a metal nodule on the wheel which clicks a microswitch on

rotation. These switches have been labelled in Figure 2.19 along with other components discussed here. The entire mask wheel assembly sits on a plate that is mounted on ball-bearings and locked down with screws which provide tip-tilt adjustment and allow the height of the wheel to be changed slightly. An XY adjustment of the whole assembly is provided by opposing push screws at the sides of the plate. Once adjusted to the correct position the wheel-plate is clamped rigidly to the top plate.

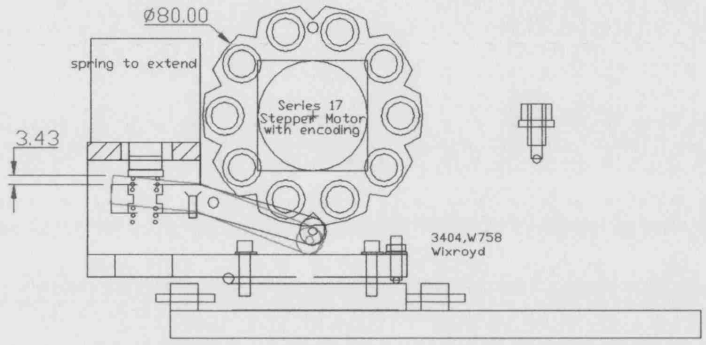

Figure 2.20: Drawing of the OSCA wheel assembly (Andy Charalambous), showing how the detent arm is moved up and down to lock the wheel in place.

Since OSCA is located at a Nasmyth focus of the telescope, the telescope field derotator (which ensures the observation object does not rotate as the telescope tracks on its alt-az mounting) has the effect of rotating the telescope aperture. Since the Lyot masks used for OSCA include vane masking, the Lyot mask must be rotated to maintain the alignment with the telescope pupil image. The Lyot mask rotation is performed by a Newport SR50 step-motor rotation stage, pictured in Figure 2.17. This has an on-axis clear aperture that allows for easy mounting of the mask on to the rotator. The SR50 is bolted to a plate that sits on three ball-tipped locking screws (as for the whole top plate). Again these sit on three pads of hardened steel - a flat, groove and cone. This means that positioning of the Lyot mask can only be achieved by tilting the entire Lyot mask assembly by altering the height of one or more of the SR50 locking screws.

All flat mirrors are seated on top of Edmund Scientific dovetail slide adjusters to provide exact placement of the mirrors along the beam. All kinematic mirror supports were commercially bought and minor alterations were made where necessary. The mirror mount

Table 2.3: Summary of opto-mechanical components and their functions

Component	Required functions	Mechanism	Accuracy
OSCA (entire assembly)	Move along optical path by ± 5 mm	Manual screw stops	$\pm 50\,\mu$m
Top plate	Deployable up/down with 125mm stroke	Pneumatic actuated against manual adjust hard-stop	$\pm 5\,\mu$m
	± 3 mm vertical and ± 1 mm rotational adjustment	Manual adjust on 3-screw mounting	$\pm 20\,\mu$m
Wheel-assembly plate	Translation in XY ± 1 mm Focus adjust 2mm	Manual adjust on 3 screws over bearings	$\pm 5\,\mu$m $\pm 40\,\mu$m
Focal plane mask wheel	10 position rotating wheel lockable in position	Home switch Stepper motor Pneumatic detent arm	$\pm 5\,\mu$m mask relocation
Lyot stop	360° mask rotation with home switch	SR50 step motor rotation stage	$\pm 0.06°$
	Translation in XYZ ± 1 mm	Manual adjust on 3-screw mounting	$\pm 40\,\mu$m
Flat mirrors & OAP2	Rotate beam about XY $\pm 2°$ Move along beam ± 2 mm (flats only)	Kinematic mirror support (2 screw) Slide adjuster	$\pm 0.045°$ $\pm 40\,\mu$m
OAP1 mirror	Rotate beam about XY $\pm 10°$ Piston ± 2 mm	Kinematic mirror support (3 screw)	$\pm 0.0042°$ $\pm 40\mu$m

units were custom designed, see Appendix B. Care had to be taken over the height of mechanisms above the top plate to ensure they were beneath the OASIS beam. The most problematic component in terms of clearance from surrounding instrumentation on the bench was OAP2: this was extremely close to the OASIS pick-off mirror. Since OAP2 has no height adjustment this mirror defined the height of the optical system above the top plate. To ensure an accurate alignment OAP1 was mounted in a 3-point adjustable kinematic mirror unit and seated on a height adjustable plate which is locked on to the top plate with screws.

All aluminium parts were black anodised to reduce stray reflections. OSCA requires a clean environment since dust increases scattering; a cover has been planned, perhaps with over-pressurised air flow, to keep OSCA dust free.

2.4 Electronic and Software Design

The electronic control systems were designed and assembled at Durham University. This was based on the preference for the controls to be similar to those of NAOMI (built by Durham University and UK ATC) and their expertise with EPICS - the software control language used at the telescope. Graphical user interfaces (GUIs) for observers were written at the ING. A summarised description of this work is included here for completeness.

OSCA mechanisms that required electronic control were:

1. The stepper motors on the mask wheel and SR50 rotation stage, including any encoder feedback and home switches.
2. The two pneumatic cylinder controls - open/close valve to pressurise or release, one to deploy OSCA into the beam and the other to lock the detent arm into the wheel.
3. Signals from two limit switches attached to the deployment pneumatic cylinder to indicate position status.
4. Signals from 3 other microswitches attached to the mask wheel assembly to indicate detent status and wheel home position.

The encoder signals from the mask wheel (home position and 2-phase output) are sent to an Oregon VME58 card: this drives the stepper motor to the correct position. The

Lyot stop home signal also goes to the Oregon card, again driving the stage to the home position and resetting the internal position counter. The pneumatics are controlled by a Xycom240 card which provides the signals to the valves on both the cylinders (detent arm and deployment stage) and detects the signals from the limit microswitches for the in/out position status. A schematic of the system is shown in Figure 2.21.

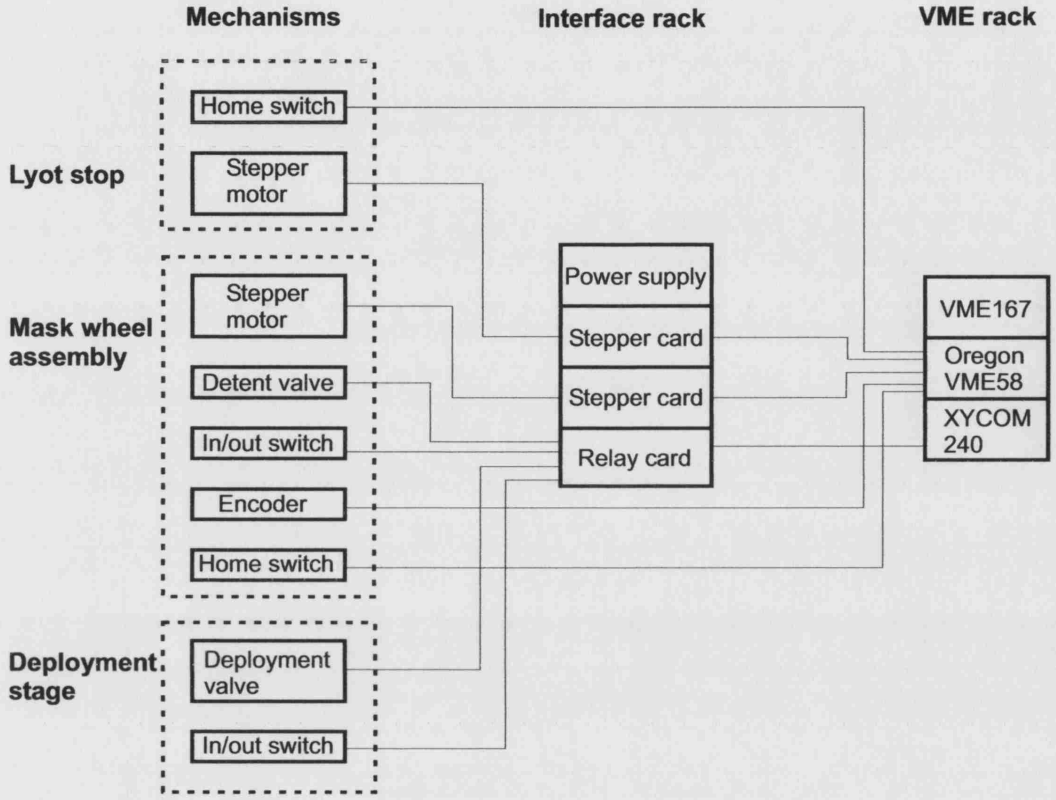

Figure 2.21: Layout schematic of the OSCA electronic system as given in the Critical Design Review documentation.

OSCA provided its own 19-inch enclosure that comprises two sub-rack assemblies, the first to house the VME167, Oregon and Xycom cards and the second a custom electronics interface rack containing the stepper motor drive cards, relays for the pneumatic values and power supply units. To simplify commissioning at the telescope this complete rack could simply be placed on top of the existing ELECTRA crate in the GHRIL control room. NAOMI uses an older version of the EPICS software which does not support the Oregon VME58 card and this was another reason a separate crate was required.

The OSCA-side wiring, which provides the panel of sockets into which the control cards

would be plugged was carried out at UCL. This consists of three 9-pin and one 25-pin D-sub type connectors. The 25-way connector comes from the SR50 step rotation stage and was unchanged from its original wiring. The pins on this were connected to the 4-phase stepper motor, home switch signals and 5V/0V power supply. The first 9-way connector was wired to the mask wheel stepper motor to provide stepping pulses and power supply and to the mask wheel home position microswitch. The second 9-way connector was wired to the encoder attached to the mask wheel stepper motor. The third 9-way connector was wired to the microswitches and pneumatic valve controls for the detent arm and deployment stage with the following pin connections:

1. Stage In Switch
2. Stage Out Switch
3. Switch 0V
4. Detent In Switch
5. Detent Out Switch
6. Deploy Valve (5V)
7. Deploy Valve (0V)
8. Detent Valve (5V)
9. Detent Valve (0V)

The valves have suppression diodes fitted to them to prevent any back-emf. To avoid any possible noise problems ground lines were put in for the valve signals and the deployment and detent switches. Additionally, all the cables were shielded by soldering the cable shield to the D-connector bodies and the connectors are linked to the OSCA metalwork which is earthed through the optical bench.

The cables were placed to ensure that there is enough clearance for OSCA to fully deploy out of the beam. The cable positioning can be seen in Figure A.5.

The control software for the OSCA mechanisms was written in EPICS to be consistent with the system already in place at the telescope. Apart from providing information as to the sequence of events that the mechanisms should undergo, UCL was not involved in the writing of this software. A brief summary of the event sequence for the main mechanisms is given here.

The software stores the position of the different occulting masks in the wheel by counting

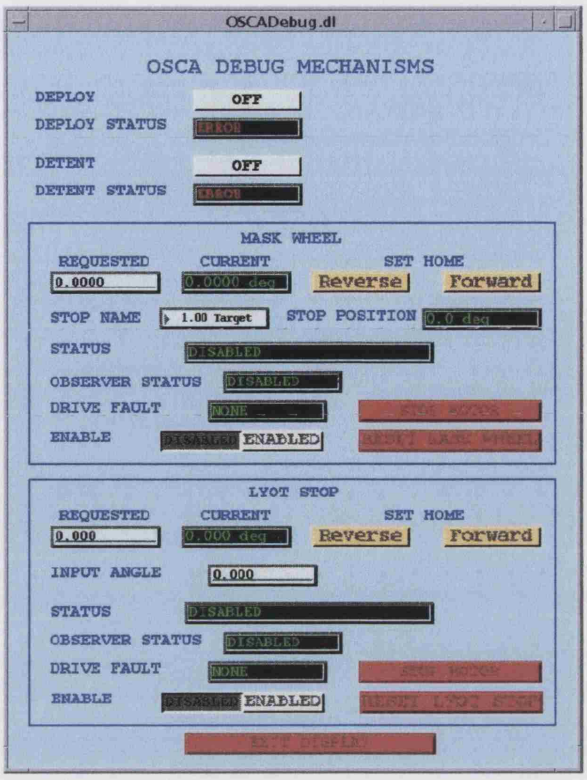

Figure 2.22: Screenshot of one page from the OSCA engineering GUI.

steps from the home position. The home position is located by driving through it a few times from either direction and the counter position stored. A look-up table provides the number of steps to each of the different masks. The sequence for positioning the mask wheel in the control software is as follows:

1. Release detent arm
2. Switch motor on
3. Step certain number of encoder pulses until near correct position
4. Switch motor off
5. Engage detent arm

The basic properties of OSCA are entered straight into the EPICS program files: these are the lists of the step positions of the home switches and different occulting masks. The next control level up are the OSCA Engineering GUIs which provide easy and flexible control of all the moving parts to allow complete testing and tracing of problems with the mechanisms. One of the engineering screens is shown in Figure 2.22.

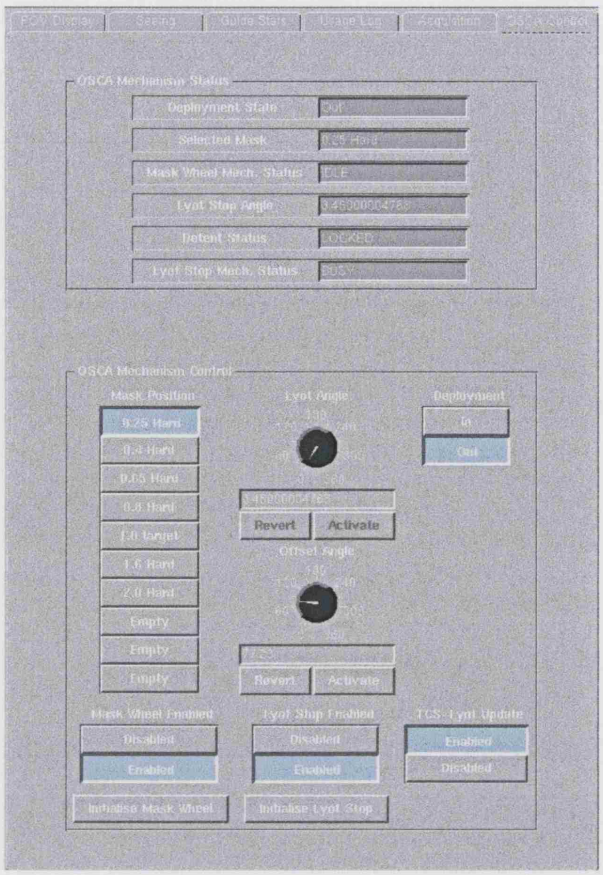

Figure 2.23: The observers GUI used for controlling OSCA.

The top level control, for use during normal observing, allows the observer to easily select the different OSCA options as shown in Figure 2.23. This was designed and implemented at the WHT by Craige Bevil and is an additional tab on the NAOMI/INGRID observers GUI.

The Lyot stop required some additional software to ensure its rotation was synchronised in time with the derotator of the telescope. The derotator angle is updated 20 times per second by the telescope control system (TCS) and is calculated from the right ascension, declination and sidereal time. The process responsible for updating the Lyot stop obtains the required information from the TCS. The alignment procedure for the Lyot stop is discussed in §4.

CHAPTER 3

Simulations

Computer simulations have been performed for a coronagraphic system with a segmented adaptive optics corrector (as in NAOMI). The main goals of the simulation were as follows:

1. To determine the optimum size for pupil plane masking over the specified wavelength range (0.4 - 2.2μm with emphasis on the near IR) and a range of focal stop sizes. To determine whether vane masking should be used and of what size.

2. To predict the system performance with the chosen pupil masking, i.e. the amount of suppression provided by the coronagraph in a variety of different configurations - focal mask sizes, wavelengths, seeing conditions (r_0) etc.

3. To investigate the effect of some simple but novel coronagraphic masks e.g. gaussian masks, square shape masks and compare their performance (like size for like) with the traditional hard-edged stops.

The principles of modelling an optical system using Fourier transform theory were summarised in §1.1. Based on this theory representative telescope apertures were described in computer code (C-code compiled on a DEC-ALPHA or a Pentium PC and MATLAB on a Pentium PC). These apertures were numerical arrays with zeros for opaque areas and ones for transparent regions. Values between 0 and 1 or complex numbers can represent partially transparent or phase shifting sections respectively. A fast Fourier transform (FFT) is performed on this aperture array to create a simulated PSF (focal plane). It is

at this stage that a coronagraphic focal-mask function is applied by multiplying it with the PSF. An inverse FFT is performed on this new array to create a simulated pupil plane in the coronagraph. The Lyot-stop function is multiplied with this pupil plane array and a final FFT is performed to produce the final image plane of the simulated coronagraphic system for a perfect flat wavefront. A sequence chart of this basic program is summarised in Figure 3.1.

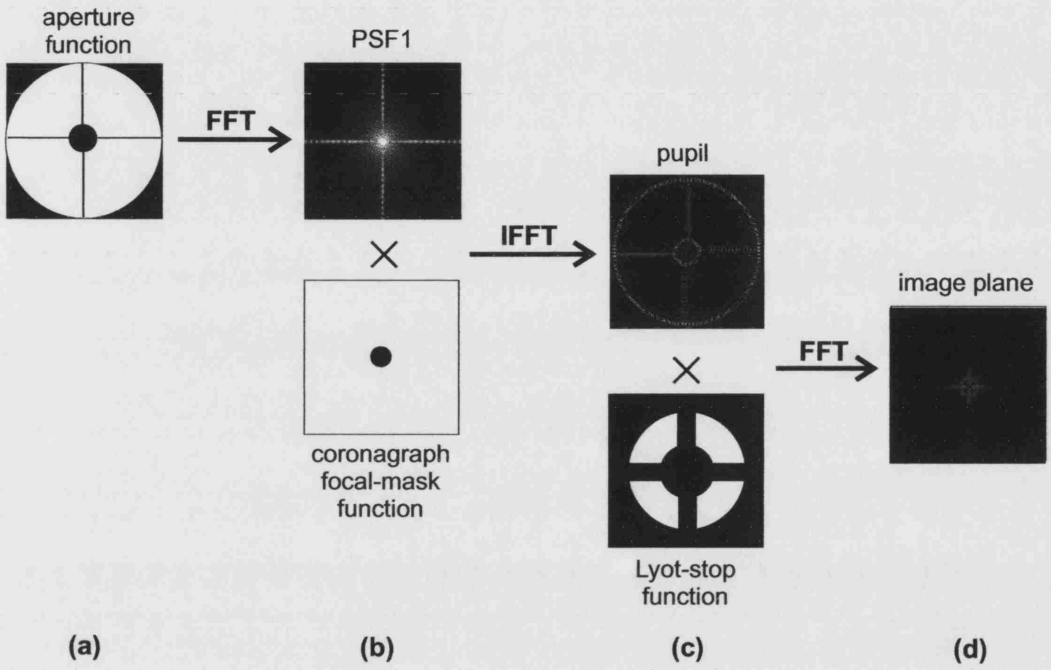

Figure 3.1: Sequence chart of basic coronagraph program simulation. The Fourier transform produces a complex array of which the modulus-squared must be taken to obtain the image-plane.

The complexity of the simulation can be developed further by feeding in a sequence of Kolmogorov-type turbulence phase screens, one of which is shown in Figure 3.2(a). These are generated using a C-program (move.c) written by A. Glindemann (Lane *et al.* 1992). The total wavefront is then represented by $\cos\alpha + i\sin\alpha$, where α is the array of phase angles generated by the Glindemann program. Each phase screen in the time sequence array-stack can be multiplied with the aperture function (Figure 3.1(a)) and processed in turn, with each subsequent final image (Figure 3.1(d)) added to create a simulated image with an integration time on the simulated sky dependent on the number of phase screens processed. This simulates what is termed here as an 'uncorrected image', i.e. one that has

not been corrected by adaptive optics.

To obtain a rough estimate of this time, the relation for the temporal coherence parameter, τ, can be used. When a beam is much larger than the coherence length r_0, τ gives the approximate time for the wind (velocity, v_w) to transport the beam one coherence length.

$$\tau \approx 0.53\,(r_0/v_w)\,(D/r_0)^{1/6} \qquad \text{(Tyson 1993)} \tag{3.1}$$

Depending on the exact parameters used in the turbulence generator and the size of the simulated aperture, this corresponds to ~ 2000 phase-screens for 5 seconds of data.

An off-axis star can be added into the simulation by adding an overall tilt to the phase array – phase further from the centre is changed by more than phase closer to the centre. It is then scaled by a chosen magnitude (brightness) difference. This off-axis star is then processed alongside the on-axis star and the two arrays added together to form the final image.

The processing to simulate an adaptive optics element is added in between stages (a) and (b) in Figure 3.1. The aperture function is multiplied by the DM mirror mask to produce the shape shown in Figure 3.2(b). The phase distortions are sensed using a Shack-Hartmann based technique and corrected using a zonal method. The zonal method divides the wavefront into N subapertures and the phase at each expressed as a number - the entire wavefront is the characteristic set of N numbers. If $N \rightarrow \infty$ the wavefront is exactly represented. For finite N, each subaperture can be represented by its fundamental modes, e.g. piston and local tilt. Currently the zonal method is the only one employed at the WHT and so is the only one used in these simulations. The wavefront is divided into an 8×8 grid (normal sensing mode in NAOMI) for these simulations. The other correction method is modal: the wavefront is described by a set of coefficients of the modes of a polynomial expansion over the entire aperture, e.g. using Zernike coefficients. This method is still under development at the telescope.

Padding of the arrays for the Fourier transform was used in the simulations. This places the array to be transformed in the centre of a larger array of zeros. The size of the array and the amount of padding determine the resolution (number of pixels across the PSF peak) in the focal plane. The resolution in the focal plane determines how well represented the shape of focal masks are (particularly important for the smallest ones). The effect of padding is illustrated in Figure 3.12. For the C-code the aperture array was

always padded out to twice its original size, the MATLAB code generally used a padding factor of 4.

Determining the size of objects in the image plane of the simulations is dependent on the size of the array used for the aperture function, its corresponding real-life size, the FFT padding factor used and the wavelength of the simulation. A resolution element (a pixel) in the image plane is therefore related to $1.22\lambda/D$, with the relation being tied to a constant factor depending on the details of the simulation set-up.

3.1 C-code simulation

The original adaptive optics C-code was written by A. P. Doel to model the ELECTRA/NAOMI system. The program was extended to include coronagraphic subroutines and renamed osca.c. The set of input parameters and example values for the osca.c program are given in Table 3.1. Apart from an optimisation to obtain the best values for gain1 and gain2, the adaptive optics model was unmodified.

The resolution of the model was increased from a 64 × 64 aperture array to 128 × 128, padded to twice its size to produce the PSF over a 256 × 256 array. A larger array size would have been preferable, so as to better represent some key aperture features (such as the telescope vanes which were added to the aperture function), however this was not possible due to computing resource limitations at the time. The AO sensor wavelength used in these models is non-standard, and is always taken to be $\lambda = 0.7\mu$m, therefore this is also the wavelength for which D/r_0 is given. Observation (science) wavelength is variable and is specified for each simulation.

As already mentioned, the input to the simulations was generated by a C-program written by Glindemann. The phase-screens simulate a single turbulence layer in the atmosphere, moving across the aperture in the y-direction to represent a constant wind velocity. A decorrelation factor (0 - no change, 20 - slight, 1000 - decorrelated) alters the phase structure between frames to model the effect of turbulent eddies etc. Raising the value for the decorrelation factor has the effect of creating more differences between subsequent frames in the sequence and so simulates a greater level of turbulence. A typical set of parameters to generate the phase-screens produced by this code is given in Table 3.2. The

Table 3.1: Input parameters and some example values for the AO-coronagraph program osca.c

Input parameter	Description	Value
numframes	Number of phase-screens to process	2000
dr0	D/r0, seeing at sensor frequency	20
mag	Guide star magnitude	7
gain1	Quad-cell gain	2
gain2	Tilt gain	0.1
hys	AO hysteresis (%)	1
sdev	RMS pixel noise	0
algo	AO correction algorithm - sor/mod	
wave	Wavelength (microns)	1.6
modeno	Number of modes corrected (for mod only)	32
conv	Convergence limit (for sor only)	0.01
focstop	Focal stop size (in units of λ/D)	5
focoffset	Centring offset (in pixels) of focal mask over PSF	0
lyotstop	% primary masking	80
lyotrad	% oversized secondary masking	20
focmode	Type of focal stop - Disk/Gaussian/Square/Fuzsquare/Gaussdisk	
PreAOmode	Option to put in a pre-AO focal stop - Yes/No	
magdiff	Magnitude difference of off-axis star	7
angdiff	Angular separation from centre to off-axis star (arcsec)	1
foctrans	% transmission of focal stop	0

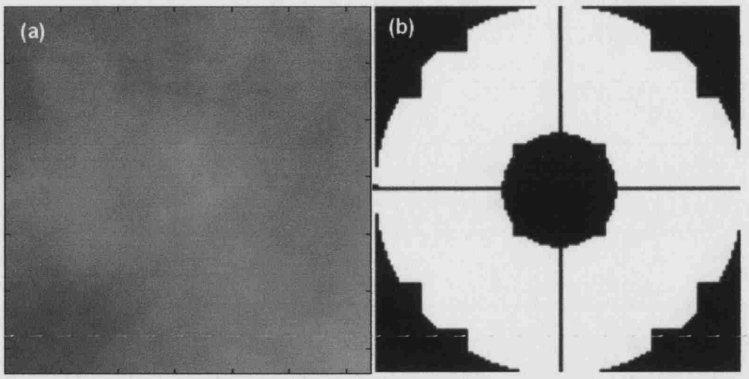

Figure 3.2: (a) One of the 2000 phase-screens generated as input to the osca.c program. (b) The aperture function, the shape is unusual because the 8×8 section of NAOMI mirror has the outer segments turned off and blacked out. Both arrays are sized 128×128.

turbulence program also has the ability to simulate the effect of a number of low frequencies that normally would not fit across the array. The size of D/r_0 for these screens was chosen to be 1, since the AO program multiplied the phase-screens by $(D/r_0)^{5/6}$, so the same set of inputs could be used for a variety of seeing conditions.

A full simulation used a set of 2000 of these Kolmolgorov-type turbulence phase-screens. The array size representing the telescope aperture is picked to match the array size of the phase-screens, this was the limiting factor in deciding a suitable aperture size. To obtain a reasonable sampling of the seeing, it was determined that a minimum of 2000 phase-screens, corresponding to approximately 5 seconds, should be processed. This limited the array size to 128×128, based on available computer data storage at the time.

The estimate for the exposure-time of the final image is calculated using Equation 3.1. Taking a value of 15ms^{-1} for the wind velocity, $D/r_0 = 20$ and $D = 4.2$ for the WHT, gives $\tau \sim 0.012$ seconds. Since there are 128 pixels across the aperture, this corresponds to 6 pixels per r_0. Given the shift between each phase-screen to be 3 pixels, a set of 2000 screens translates to approximately 10 seconds. An example of one of the phase-screens generated by move.c is shown in Figure 3.2(a).

The AO code had the choice of using a zonal (sor) or modal (mod) correction. The zonal option was always used since this is what is implemented at the telescope. The zonal method also provides better correction when using bright targets as is usually the case in

Table 3.2: Input parameters and the values chosen to generate the turbulence phase-screens using move.c

Input parameter	Value
Size of array (pixels)	128
Relation size of phase-screen to size of array	2
Relation size of support to size of array	1
D/r_0	1.00
Number of low frequencies	5
Number of phase-screens	2000
Number of photo events	0
Shift of phase-screen (pixels)	3
Decorrelation factor	20.0

coronagraphic observations. The level of hysteresis of the AO simulation was tested at 2% (as predicted for NAOMI) and at 0%, with no measurable difference. It is likely that the effect of hysteresis would not become apparent over such short timescale simulations. The natural guide star magnitude was set to 11 for the majority of the simulations, however a guide star magnitude of 7 was also tried and again no measurable effect on performance was noticed. A variety of seeing conditions were tested (D/r_0 from 5 - 20); the AO cannot correct for values of D/r_0 higher than 30 and can actually worsen the image quality. Since $r_0 \propto \lambda^{6/5}$, D/r_0 decreases (the seeing improves) with increasing wavelength; the quad-cell gain was scaled to account for this, a necessity since the sensor operates at one wavelength (0.7μm) while the mirror must be set to correct at another wavelength. The simulations were performed at monochromatic wavelengths but in reality the system operates over a range of wavelengths dependent on the filter used. This would have the effect of blending monochromatic image features over a wider area (diffraction rings etc.) by an amount proportional to the wavelength range.

Simulation options for the coronagraph were the size and shape of the focal mask, the size of the pupil mask and an option to include a pre-AO focal mask. If the pre-AO stop was selected, its size would be set to the same size as the focal mask.

At this time the exact configuration of NAOMI was uncertain. The simulation took an

8×8 segment mirror (a down-sized NAOMI mirror) to fit across the array. The telescope primary aperture was applied to this so that all segments lay within it, hence the unusual shape of the aperture function in the simulations (Figure 3.2(b)).

3.1.1 Determining the Lyot stop size

A set of computer simulated experiments were performed to determine a suitable size for the pupil plane mask. The effects of the percentage primary and secondary mirror masking, wavelength and focal stop size upon the amount of suppression were investigated. Since the original AO code did not simulate the effect of the telescope vanes, these were added to the aperture function so that it was possible to determine the need for vane masking in the Lyot stop. Due to the low resolution of the aperture array, the vane width was represented by 1 pixel. The vane offset from the centre was not included in these simulations.

The size of the Lyot stop is specified as a percentage undersized from the primary mirror and a percentage oversized from the secondary. As described in §1.2, this is to mask out the over-intensity of light around the edges of the aperture caused by the addition of a stop, acting as a spatial filter, in the focal plane prior to this pupil plane. The size of the focal stop alters the amount of light removed from the PSF and thus changes the distribution of light in the pupil plane. The wavelength used also has a similar effect since for a fixed sized mask (e.g. 1 arcsecond), the number of Airy rings covered by the mask increases with decreasing wavelength. To visualise the change in light distribution in the pupil plane due to these various parameters, consider Figure 3.3.

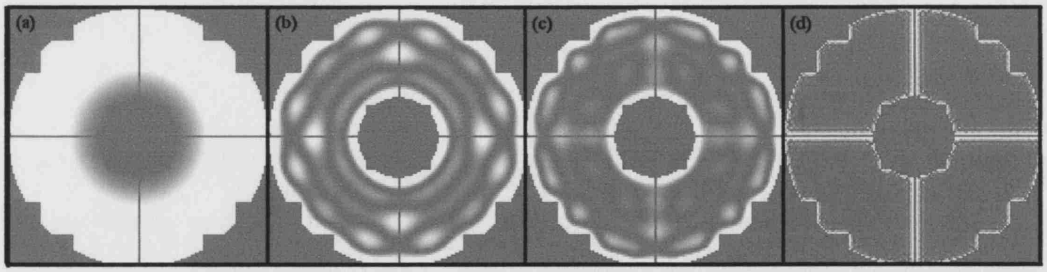

Figure 3.3: Simulated pupil images (flat wavefront) at (a) $\lambda = 2.2\mu$m, focal-stop=0.2", (b) $\lambda = 0.4\mu$m, focal-stop=0.2", (c) $\lambda = 2.2\mu$m, focal-stop=2.0" and (d) $\lambda = 0.4\mu$m, focal-stop=2.0"

The simulated pupils in Figure 3.3 demonstrate the effect of wavelength (0.4 - 2.2 μm) and focal-stop size (0.2 - 2.0 arcsec) on the pupil plane in OSCA. From this it can be inferred that a small (reduced throughput) Lyot stop is needed to give the best suppression when using smaller focal masks and longer wavelengths, whereas a larger Lyot stop is adequate in removing the over-intensity when larger masks or shorter wavelengths are used. In these pupil images the aperture mask is overlaid for reference, however the vanes under this are bright (shown in Figure 3.20).

Sets of simulations were run for different sized Lyot stops (no vane masking at this stage) with a range of focal mask sizes at a wavelength of 2.2μm. The longest wavelength was used since the focal masks have the smallest apparent size at this wavelength and so the suppression performance has a strong dependence on the amount of pupil masking. For a 1 arcsecond focal mask, the amount of suppression obtained with a range of different Lyot stop sizes is given in Table 3.3.

Table 3.3: A selection of Lyot stop configurations for the primary and secondary masking, with the suppression obtained. Simulations were at $\lambda = 2.2\mu$m, using a 1.0 arcsecond focal stop. The suppression was measured 0.5 arcseconds from the focal mask edge.

Lyot stop configuration (%)	Throughput (%)	Suppression factor
100,0	100	0.20
100,30	93	0.39
90,10	77	0.43
90,20	74	0.43
90,30	72	0.43
80,0	60	0.57
80,10	58	0.57
80,20	56	0.61
80,30	53	0.64
70,30	36	0.82

Choosing a suitably sized Lyot stop is a compromise between system throughput, image resolution (the Lyot stop undersizes the aperture, so effective D is smaller) and the sup-

pression performance of the coronagraph. Taking these factors into consideration, a Lyot stop size of 80% primary masking and 120% secondary masking was chosen for OSCA. For increased system throughput an additional Lyot mask of 90%, 10% was also chosen. At shorter wavelengths ($\leq 1.2\mu$m) and when using the larger focal masks, the amount of suppression is similar to that of the smaller Lyot stop (accounting for the throughput differences).

Masking for the vanes of the telescope was then added to the Lyot stop function to determine the effect on the amount of suppression. Since the width of the vanes in the aperture was represented by one pixel (130% of the actual size), the width of the vane masking was trialled at 3, 5, 7 and 9 pixels. As for the edges of the aperture, the spread of over-intensity around the vanes in the pupil changes with the size of focal masks used and the wavelength of the simulation. Vane masking at 3 times the vane width provided very little gain in the suppression performance. However there was a 5% improvement on suppression performance on using 5 times vane-width masking compared to no vane-masking (accounting for change in throughput). For larger vane masking, it was decided that the small gain in suppression performance was not worth the associated loss in system throughput and PSF quality.

3.1.2 Coronagraph suppression results

To provide an estimate of expected system performance, simulations were run for a range of focal mask sizes at different wavelengths. The size of pupil masking was kept fixed at 80%, 20%, since this would be the usual size of Lyot stop used on OSCA.

The intensity pattern of diffraction by a circular aperture is an Airy disk that is characterised by a first order Bessel function, $I \propto |J_1(x)/x|^2$. For an aperture of width D and a wavelength of light λ, the zeroes on this Airy disk occur at 1.22, 2.23, 3.34 and 4.24 λ/D. As described in §2, the size of the smallest mask was chosen to coincide with the first minima of the Airy disk at the best AO corrected and longest wavelength of 2.2μm. The sizes of the of the remaining masks were chosen in a range of steps of $1.22\lambda/D$ at 2.2μm and to ensure a good range of sizes over the entire wavelength range. The plots in Figures 3.4 - 3.6 are suppression profiles over the wavelength range 0.7 - 2.2μm for a 1 arcsecond diameter disk focal mask. Comparisons are made of the suppression when using a coron-

agraphic system with and without adaptive optics correction. In all these cases the use of AO improves the suppression near the mask, however further away there is a cross-over for the suppression curve without AO. This effect is more pronounced at shorter wavelengths, where the AO is only partially correcting the wavefront and the remaining light (of higher order aberrations) remains in a broad 'seeing halo' or is scattered to contribute to higher background light levels. The use of a focal stop before the AO system (Figure 3.4(d)) reduces light entering the system, reducing the background scattered light and bringing the suppression performance of an AO coronagraph at large distances from the mask to that of an uncorrected coronagraph. Simulations were attempted for wavelengths shorter than 0.7μm, however the AO acted to make the PSF worse than the uncorrected PSF, so the limit of useful AO correction for this particular system is taken at $\lambda = 0.7\mu$m. A selection of suppression results at different wavelengths and using different focal masks is given in Table 3.4.

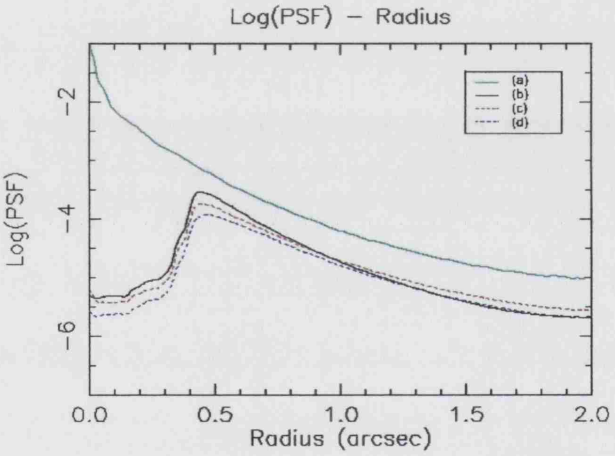

Figure 3.4: Simulated PSF profiles with $\lambda = 0.7\mu$m, f-stop=1.0", Lyotstop(%) = 80,20. (a) corrected, no coronagraph; (b) uncorrected, with coronagraph; (c) corrected with coronagraph; (d) corrected with coronagraph and pre-AO stop.

For focal masks smaller than the natural seeing disk size and at longer wavelengths where the AO can provide good correction (as in Figure 3.7), the use of AO with a coronagraph out-performs a non-AO coronagraph at all distances from the mask. Hence the provision of small focal masks in OSCA ($\leq$ 0.5arcsec in diameter) realises the maximum benefit of

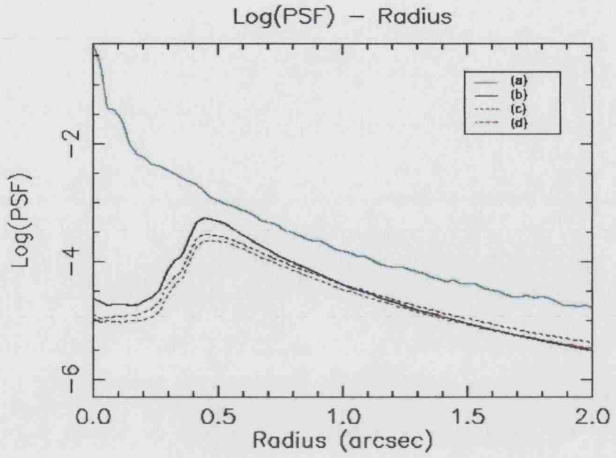

Figure 3.5: Simulated PSF profiles with $\lambda = 1.2\mu$m, f-stop=1.0", Lyotstop(%) = 80,20. (a) corrected, no coronagraph; (b) uncorrected, with coronagraph; (c) corrected with coronagraph; (d) corrected with coronagraph and pre-AO stop.

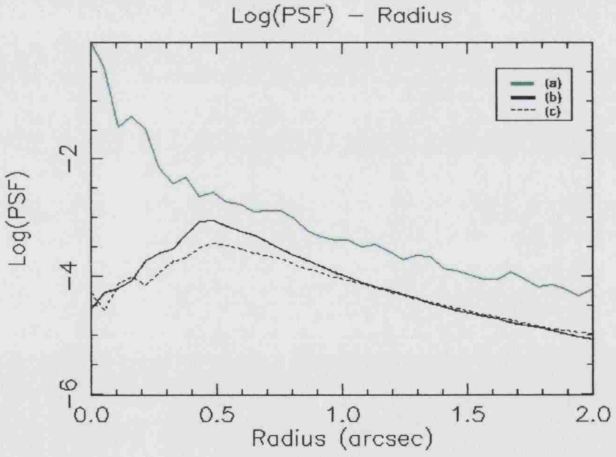

Figure 3.6: Simulated PSF profiles with $\lambda = 2.2\mu$m, f-stop=1.0", Lyotstop(%) = 80,20. (a) corrected, no coronagraph; (b) uncorrected, with coronagraph; (c) corrected with coronagraph.

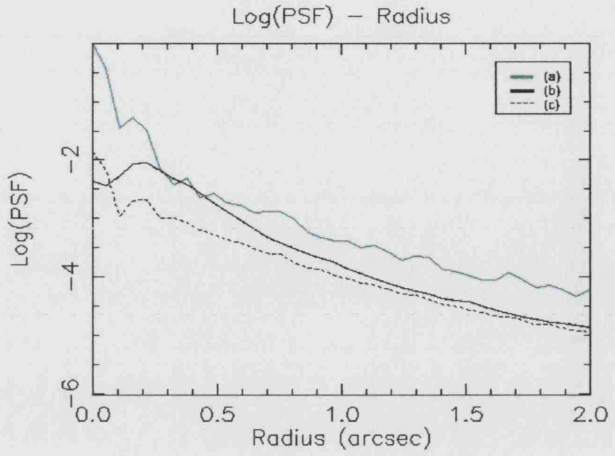

Figure 3.7: Simulated PSF profiles with $\lambda = 2.2\mu$m, f-stop=0.36", Lyotstop(%) = 80,20. (a) corrected, no coronagraph; (b) uncorrected, with coronagraph; (c) corrected with coronagraph.

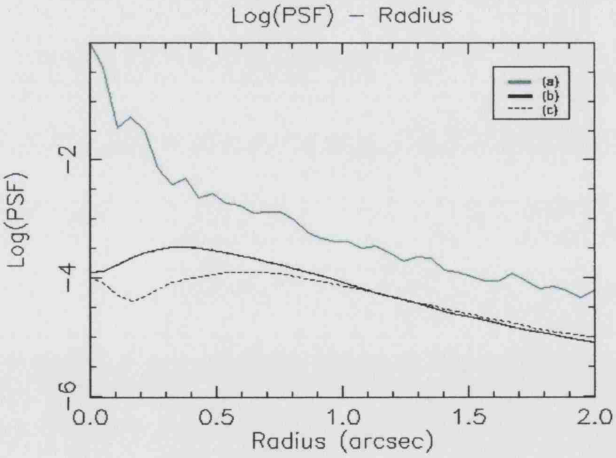

Figure 3.8: Simulated PSF profiles with $\lambda = 2.2\mu$m, gaussian shaped focal mask (FWHM = 1.0"), Lyotstop(%) = 80,20. (a) corrected, no coronagraph; (b) uncorrected, with coronagraph; (c) corrected with coronagraph.

having an AO coronagraph.

Table 3.4: Coronagraph suppression results using the osca.c simulation. Suppression results are compared to the corrected, no coronagraph PSF at 0.5 arcsec from the edge of the focal mask used. Simulations were for 2000, 128 × 128 phase-screens, at a $D/r_0 = 20$. The Lyot stop was sized at 80,20. Where use of a pre-AO stop is indicated, this is set to same size as focal mask.

Focal mask diameter (arcsec)	Pre-AO stop	Wavelength (μm)	Suppression (Δm)
1.0	No	0.7	1.43
1.0	Yes	0.7	1.79
2.0	No	0.7	1.50
2.0	Yes	0.7	1.88
0.5	No	1.2	1.13
0.5	Yes	1.2	1.35
1.0	No	1.2	1.61
1.0	Yes	1.2	1.79
0.3	No	2.2	1.75
0.5	No	2.2	1.75
1.0	No	2.2	2.14
Gaussian, 1.0	No	2.2	3.4

To illustrate the gain in using an AO coronagraph to detect a faint off-axis object, consider the simulated images in Figures 3.9 and 3.10. Even where there has only been a partial correction at $\lambda = 0.7\mu$m, the combination of AO and coronagraph enables the detection of a $\Delta m = 7$ off-axis star; this star is not apparent in the non-corrected case. The advantage of using a Gaussian shaped focal stop compared to a hard-edged disk mask is evident from the simulated images in Figure 3.11. The tapered Gaussian profile does not introduce any additional diffraction effects as the hard disk does, so there are no image artefacts to be confused with the real off-axis star. Although the off-axis object suffers some attenuation (the amount depending on the distance from the centre of the Gaussian) when using broad tapered masks such as these, the suppression performance (see Table 3.4) is significantly more than that for a similar sized disk mask and the reduction in image artefacts facilitates

the detection of faint point sources.

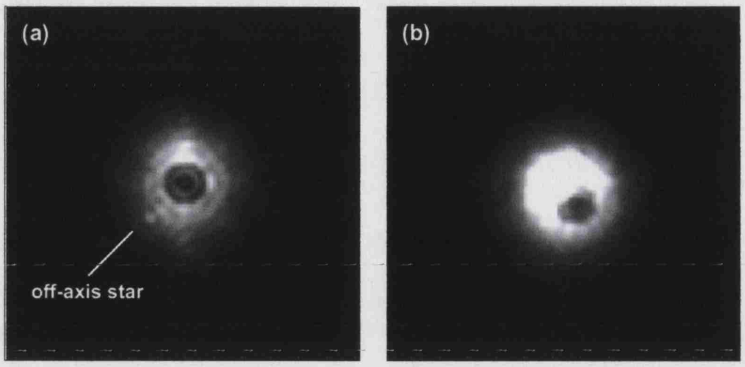

Figure 3.9: Simulated coronagraph images at $\lambda = 2.2\mu$m, f-stop=0.8 arcsec, D/r_0 = 20 for 2000 wavefronts, mag difference = 7 (a) with and (b) without adaptive optics.

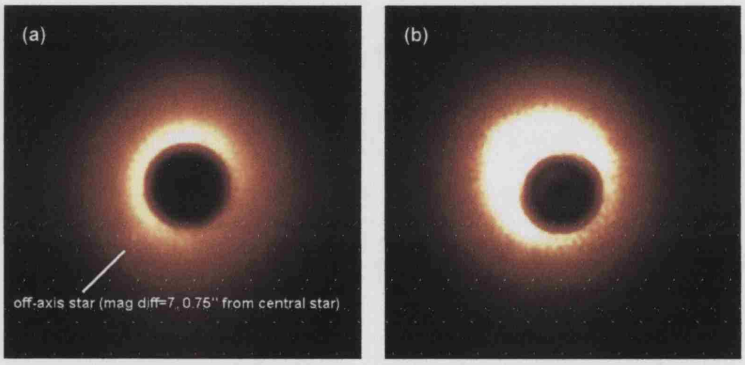

Figure 3.10: Simulated coronagraph images at $\lambda = 0.7\mu$m, f-stop=1.0 arcsec, (a) with and (b) without adaptive optics.

3.1.3 Discrepancies of C-code simulation compared to real system

These initial simulations were performed over a short time period (~3 months) at the start of the OSCA project. At this time NAOMI had only just been commissioned and precise details of instrument set-ups were not available. The pupil imager in INGRID was not available until the end of 2002, so no pupil plane images were available to assist in design. Subsequently there were found to be a number of differences between the simulated system and reality.

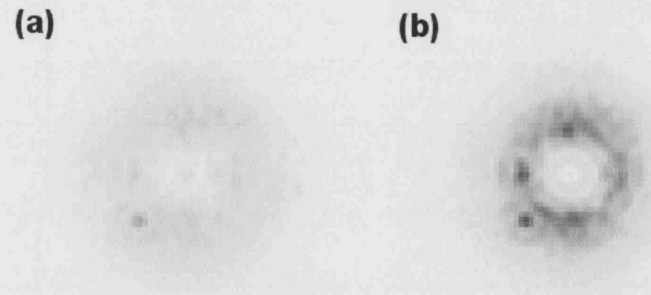

Figure 3.11: Simulated coronagraph images with AO correction and an off-axis star at $\lambda = 2.2\mu$m, using an 80,20 Lyot stop. (a) Gaussian shaped mask, FWHM = 1.0 arcsec, (b) disk shaped mask, diameter = 1.0 arcsec. Both images are clipped to same level and inverted for clarity.

The adaptive optics simulation uses 8 mirror segments across the aperture diameter while in fact only 7.3 of the NAOMI segments lie across the WHT pupil. The consequence of this is that the real system has a slightly lower order of correction than that which is simulated here. However, comparisons of PSF profiles with and without a coronagraph should be comparable with those on the real system, particularly if PSFs with similar Strehl ratios as those achieved on NAOMI are used.

During the OSCA commissioning it was discovered that 4 of the correcting mirror segments ($\sim$10% of those used in the aperture), 2 of which lie completely within the OSCA pupil mask, are faulty. These degrade the performance of both the adaptive correction and the suppression ability of the coronagraph. The faulty segments tend to piston out of phase with the rest of the mirror to create additional step functions in the aperture. Therefore the efficiency of the Lyot stop in removing the excess on-axis starlight, concentrated at the edges in the aperture, is reduced, since some of this light is conditioned around the edges of the faulty segments.

The hysteresis in the NAOMI system (long term stability) is worse than simulated. The effect of this is to degrade the PSF over time (reduction in Strehl ratio) and so reduce suppression performance of the coronagraph (less light is removed at the focal plane). Since the longest simulations were only equivalent to $\sim$ 10s, these effects would not have been apparent.

The size of the array in the osca.c simulations does not accurately represent the smallest features in the aperture, in particular the gap between the segments in the adaptive mirror and the width of the telescope vanes.

3.2 MATLAB simulation

Some time after the commissioning of OSCA, better computing facilities became available and the osca.c code was rewritten in MATLAB with the ability to use larger arrays. The phase-screen generation program was also modified so that it could be compiled on a PC and some of the original DEC-ALPHA libraries were replaced for PC-compatible ones which had been greatly optimised for speed. MATLAB has the functionality to interface with C-programs and this was used to generate the phase-screens one-by-one as needed during the AO-coronagraph simulation. This compares with the previous set-up where the phase-screens had to be generated in advance and stored as one large file. For larger phase-screens, as were required for the MATLAB simulation, this method would have been problematic. The maximum file size that can generally be indexed on a computer system is 2^{32} bytes, i.e. a 4GB file and the same size as a set of 2000 512×512 phase-screens. So the new simulation has greater scope for expandability in terms of size and number of arrays.

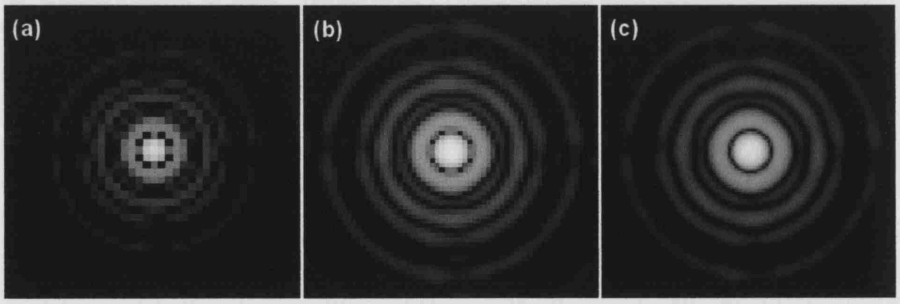

Figure 3.12: Different sampling of the PSF dependent on aperture size and FFT padding. (a) 128×128 with $2\times$ padding, (b) 256×256 with $4\times$ padding, (c) 256×256 with $8\times$ padding. The values are shown at a log scale to visualise the full range of structure in the PSF.

Aside from the requirement of a larger aperture array to better represent the smaller features, the size and FFT padding of the array is critical to obtain a reasonable sampling

of the PSF. To illustrate the importance of this, a number of PSFs were simulated using the WHT-shaped aperture, a single flat-wave input and different aperture array sizes and FFT padding factors. The results are shown in Figure 3.12. The first of these (a), represents the PSF obtained using the aperture array size and padding factor that were used for the osca.c simulations, (b) is the normal setup for the MATLAB simulations and (c) represents a higher sampling level that could be obtained if computing power and time were not an issue. The differences between these simulated PSFs is partly due to the size of a pixel (in real units, m) in the aperture array determining the lowest spatial frequency in the transform plane, while the relative size of the aperture array to the padded array determines the smallest spatial frequency, i.e. the pixel resolution in the focal plane. Due to the restriction on the size of the arrays in the osca.c program, the shapes of the smallest disk masks and Gaussian function masks were not well resolved. The MATLAB simulations, using higher resolution arrays, re-investigated these masks and some other different shaped masks.

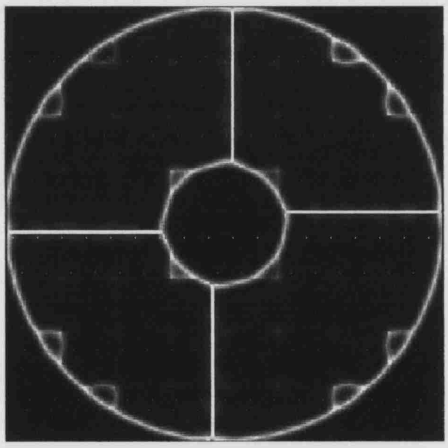

Figure 3.13: Pupil image for a full AO simulation with coronagraph (2.0 arcsec focal stop and 80,20 Lyot stop). No segment gaps were represented here, however the segmented mirror edges can be seen, especially on those were there was no active mirror correction (outer 8 and inner 4) due to phase mismatching.

In addition the MATLAB simulation altered the segmented mirror mask – in the original C-program the outer segments that were not used for correction were blacked out (set to zero) – the outer segments are not masked out and are not actively correcting the wavefront. This gives rise to pupil images as shown in Figure 3.13 and closely resembles

pupil images seen in the real system where large phase mis-matches between segments cause an over-intensity around the segment.

Unless otherwise stated the AO-coronagraph MATLAB simulations were based on a 256 × 256 aperture array, padded to 4 times its size (1024 × 1024) for the FFTs. This aperture size and padding gives reasonable sampling of the PSF (as in Figure 3.12(b)) compared to smaller padding factors. Arrays larger than 1024 × 1024 took considerable time to transform and were not used. As an example, a typical AO-coronagraph simulation with 2000 phase-screens at the fore-mentioned sizes takes approximately 4 hours to complete. To investigate the effect of some small aperture features, single-screen, flat-wave input simulations were performed using larger arrays.

3.2.1 Different shaped coronagraphic masks

Although some of these were included in the osca.c code, the resolution (sampling) of the different masks was poor. These performance comparisons were re-run in the MATLAB simulation for much larger array sizes (256 aperture embedded in 1024 grid) and thus better sampling of the coronagraph functions.

The choice of focal mask types are detailed in Table 3.5. The PSF suppression profiles of these different shaped focal-masks are plotted in Figure 3.14.

It can be seen from Figure 3.14 that the best suppression close in to the mask (0.5-1.0 arcsec) is achieved by the Gaussian, Lorentzian and Gaussdisk masks respectively. At a greater distance from the masks (> 2 arcsec), the disk mask with a pre-AO stop gives the best suppression. The segmented adaptive mirror contributes a significant amount to the diffracted/scattered light in the image, so reducing the unwanted light falling on it acts to reduce the background. The effect of smoothing the edges of sharp-edged masks improves the suppression by reducing additional diffracted light. The diamond shaped mask did not perform as well as expected – in NAOMI images, both real and simulated, it was noted that the shape of the PSF was diamond shaped due to the combined effect of a circular aperture convoluted with the square segmented mirror – a mask chosen to resemble this shape could have been advantageous over a circular disk. In areas of low contrast (in the direction of the diffraction spikes from the square segments), the diamond shaped focal mask is wider compared to the high contrast directions. It is possible the full benefit of

Table 3.5: Description of the focal-mask types in the coronagraph simulation. A characteristic width has been defined for the masks as w. For radially symmetric masks, r is defined as $(x^2+y^2)^{1/2}$. FIR denotes the use of a finite impulse response filter with the smoothing width used. Initial array is $f(r) = 1$ everywhere.

Focal mask		Function description
Disk		$f(r) = 0$ for $r \leq w/2$
Gaussian		$f(r) = 1 - e^{-r^2/2w^2}$
Gauss-disk		$f(r) = 0$ for $r \leq w/4$ and $f(r) = 1 - e^{-(r-w/4)^2/w^2}$ for $r > w/4$
Lorentzian		$f(r) = 1 - e^{-r/c}$ for $c = -w/2\log(0.5)$ and $f(r) = 1$ for $f(r)_{\text{lorentz}} > 0.95$
Square		$f(x, y) = 0$ for $\lvert x, y\rvert \leq w/2$
Fuz-square		$f(x, y)_{\text{square}}$ with FIR$(w/4)$
Diamond		$f(x, y)_{\text{square}}$ with 45° rotation
Fuz-diamond		$f(x, y)_{\text{diamond}}$ with FIR$(w/4)$

such a mask was not realised in these results since the gap between the segmented mirrors is not used for a full AO simulation – the array size needs to be very large and computing time increases by an enormous amount. Some flat-wave, coronagraph only simulations could be performed with a large aperture and the segment gaps included to investigate this. A cross-hair type mask with a central disk or gaussian would also be interesting to investigate with a segment-gap simulation.

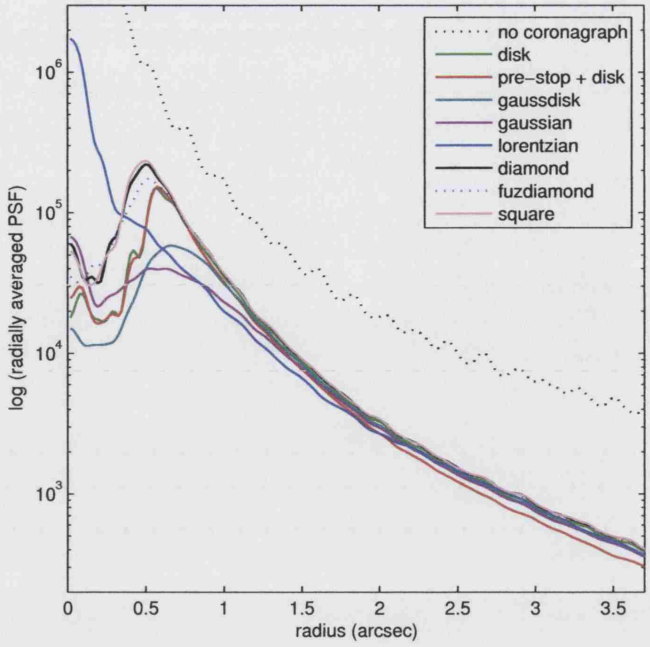

Figure 3.14: Simulated PSF profiles for the different shaped coronagraphic focal masks, the standard (80,20) Lyotstop was used for each. Characteristic width for each mask is 1.0 arcsec. Phase-screens corrected with a NAOMI-like simulated system - 2000, 256×256 phase-screens with an 4 times FFT padding factor. Peak for no-coronagraph lies at 7×10^8.

3.2.2 Effects of seeing and Strehl ratio on coronagraph performance

All of the osca.c simulations were run with phase-screens at moderate seeing conditions of $D/r_0 = 20$, producing a corrected Strehl ratio of ~ 0.35 at a wavelength of 1.2μm. For these higher resolution MATLAB simulations, at similar settings the corrected Strehl ratio is ~ 0.08. An explanation for this is that the osca.c simulations did not have enough pixels across the corrected peak of the PSF and so obtained falsely high measurements of maximum counts. It is also possible that the Glindemann phase-screen generation program is doing something unexpected when the number of pixels per r_0 increases. This was investigated to some extent but there were not any apparent errors. A more detailed study of the phase-screens would need to be undertaken to fully understand the effect which is beyond the scope of this study. To quantify the effect of the phase-screens, simulations were run at a range of D/r_0 values and the Strehl ratio for both an uncorrected and

AO-corrected PSF were recorded and listed in Table 3.6. The results for the uncorrected PSF are an optimistic estimation of the Strehl ratio for the given seeing conditions – the limited number of phase-screens results in an under sampling of the complete seeing, so that the PSF is not as broad as it really would be.

Table 3.6: Effect of seeing/AO corrected Strehl ratio on the suppression performance of a simulated coronagraph. The factor drop in suppression is normalised to the data at a $D/r_0 = 1$ and measured at a radius of 1 arcsecond from the centre of a 1 arcsecond disk mask. Seeing is given at the sensor wavelength of 0.7μm, the corrected PSF is at a wavelength of 1.6μm.

D/r_0	Uncorrected Strehl ratio	Corrected Strehl ratio	Factor drop in suppression
1	0.872	0.979	1.00
5	0.279	0.698	7.04
10	0.111	0.321	20.60
15	0.053	0.159	40.09
20	0.031	0.082	66.10
30	0.019	0.0277	148.43

The suppression performance of OSCA is strongly affected by the Strehl ratio of the PSF produced by NAOMI, as is discussed in §4. Altering the D/r_0 value was the easiest way to change the Strehl ratio to investigate the effect on the suppression. Figure 3.15 shows the radial suppression profiles for a 1 arcsecond focal disk and a 80,20 Lyot stop at a range of seeing/Strehl ratios. The factor difference in suppression for each of these, measured 0.5 arcseconds from the edge of the disk are given in Table 3.6.

To investigate the effect of mask size on suppression performance in different seeing conditions, simulations were run to cover the entire range of mask sizes in OSCA (0.25 - 2.0 arcsec), at two different seeing conditions, $D/r_0 = 10$ and 20. The results of these are given in Figures 3.16 and 3.17. For excellent seeing conditions and good AO correction the different sized masks suppress the PSF by very similar amounts. The factor difference in suppression between using the 2.0 arcsec mask to the 0.25 arcsec mask is 1.4, at $D/r_0 = 10$. The factor between the no-coronagraph PSF and the 2.0 arcsec mask is 4.6. For the moderate seeing case of $D/r_0 = 20$ the suppression difference between the largest

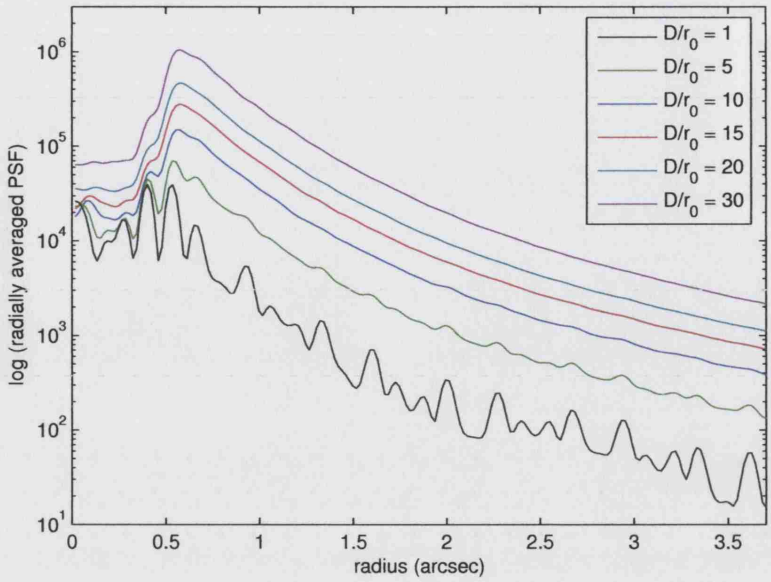

Figure 3.15: Coronagraph suppression profiles at different seeing conditions using a 1.0 arcsec disk-mask.

and smallest mask is 1.6, while the factor difference between the no-coronagraph PSF and the largest mask is only 3.2. So as expected, as the seeing degrades so does the overall suppression performance of the coronagraph and the performance difference between using the smallest and the largest mask increases.

3.2.3 Effect of mask transmission on performance

The focal plane masks in OSCA were discovered to transmit more light at longer wavelengths. To model the effect this might have on the suppression performance, a set of simulations were performed using a 1 arcsecond disk mask with transmissions of 0, 0.1, 1 and 10%. These transmission levels are equivalent to ND values of ∞, 3, 2 and 1. A 2000 phase-screen AO simulation with the coronagraph using an 80,20 Lyot stop was run for each of these disk transmissions and the results plotted in Figure 3.18.

The suppression factors (no coronagraph/with coronagraph) for these masks, starting with the 0% transmission are: 4.80, 4.80, 4.65 and 3.83 respectively. Therefore for small amounts of transmission, 1% and less, there is negligible loss of suppression performance compared to a totally opaque mask. For a 10% transmission mask the difference becomes

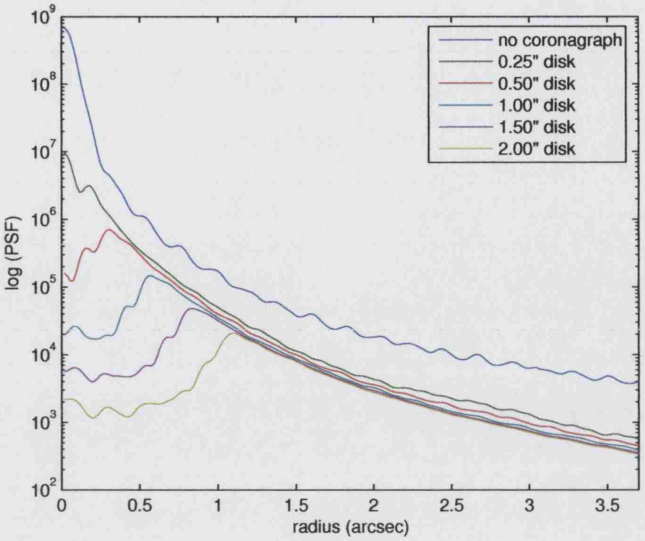

Figure 3.16: Coronagraph suppression results for a set of different sized disk masks at a $D/r_0 = 10$.

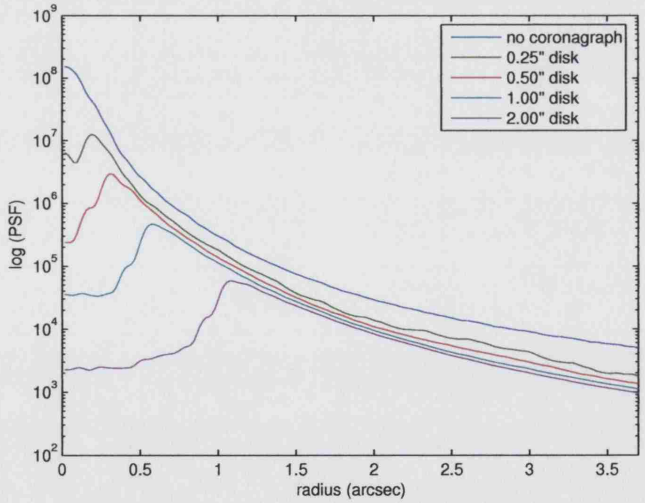

Figure 3.17: Coronagraph suppression results for a set of different sized disk masks at a $D/r_0 = 20$.

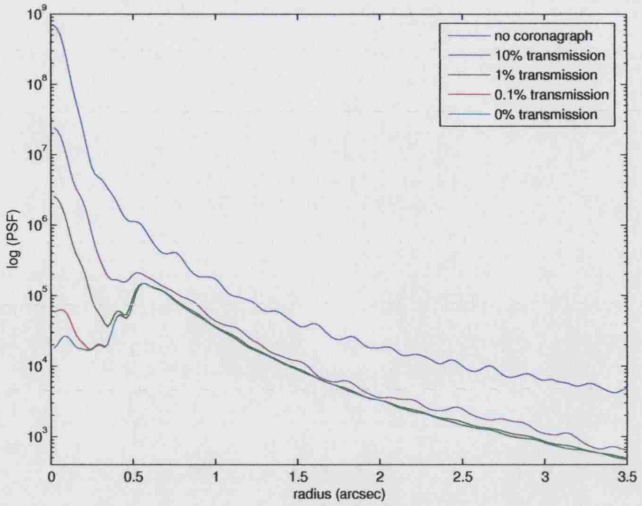

Figure 3.18: Change in coronagraph suppression performance for disk masks 1 arcsec in diameter with different transmission.

significant, with the factor loss in suppression measured at 1 arcsecond radius being 1.25 compared to the opaque mask.

3.2.4 WHT aperture features

The aperture simulated in this case was based on actual INGRID images of the pupil and recent NAOMI information (number of segments across the WHT primary). The minimum aperture size which can represent the details found in the aperture image is 512×512, which is padded to 1024×1024 to reduce the Fourier transform effects. This allows the smallest detail (the gap between DM segments) to be reasonably dimensioned to 1 pixel. The width of the telescope vanes can also be better represented using 3 pixels instead of the 1 pixel in the osca.c simulations and the vane offset pattern is also included. Figure 3.19 shows two such simulated apertures, one being the WHT pupil and the other with a segmented mirror pattern imposed. The gap between the NAOMI segments is estimated to be 0.1mm, which converts to a one pixel gap in these apertures.

The PSFs produced from the apertures in Figure 3.19 were compared and the ratio of Strehls (PSF(b)/PSF(a)) is found to be 0.97. Therefore the gaps in the segmented mirror contribute to a 3% loss in Strehl compared to a mirror without gaps.

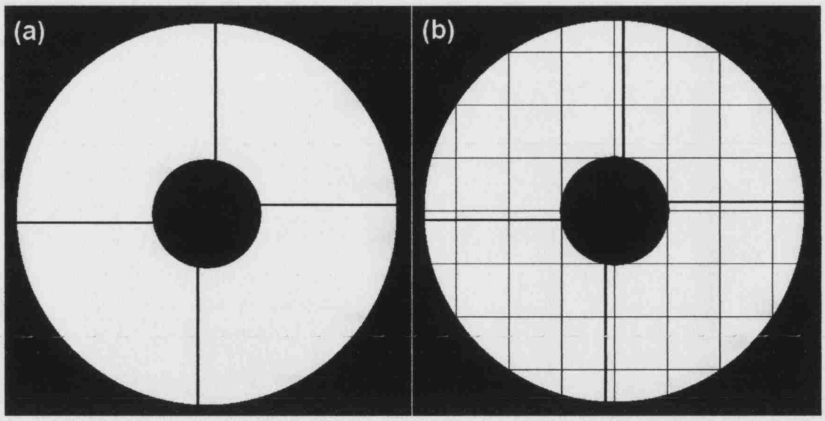

Figure 3.19: (a) Simulated WHT aperture. (b) Simulated WHT aperture with segmented mirror grid, gap size is 1 pixel ($\sim$ 0.1mm). Aperture is 512 pixels wide, embedded in an array size 1024×1024 for FFT.

The effect of the gaps on coronagraph performance, as for the faulty segments, is to detract light away from the areas normally masked by the Lyot stop. Figure 3.20 shows the pupil plane after a Gaussian focal mask for the apertures given in Figure 3.19. The simulation is for a flat wavefront, and the gaps are only represented by a one pixel width, however it indicates that large amounts of light in the pupil are concentrated around the segment edges. For the pupil in Figure 3.20(b), 45% of the total light in the pupil is concentrated around the segments. The use of a 80,20 Lyot stop with vane masking reduces the amount of light remaining in the pupil to 24% of the total amount before Lyot masking, 94% of

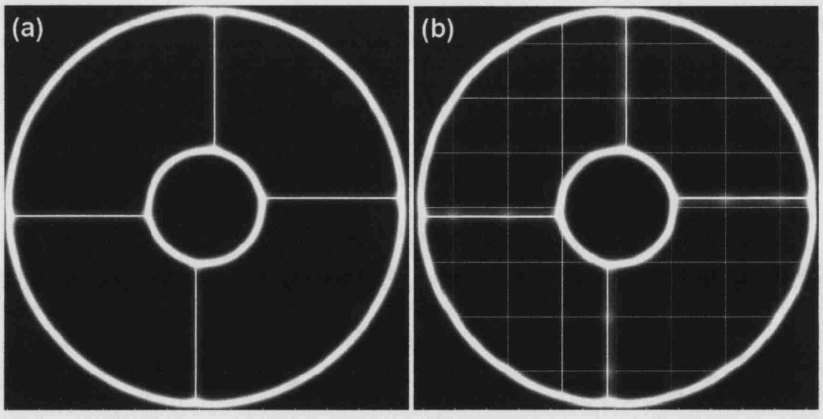

Figure 3.20: Simulated pupil images of the apertures in Figure 3.19(a) and (b), after the insertion of a Gaussian focal plane mask.

this remainder is around the segments. The resultant PSFs are shown in Figure 3.21. Avoiding the bright central diffraction spikes, the background level (several arcseconds from occulting mask) is measured to be 15 times higher in the simulated image that used an aperture with segment gaps and closer to the occulting mask the level is approximately 3 times higher than a non-segmented aperture.

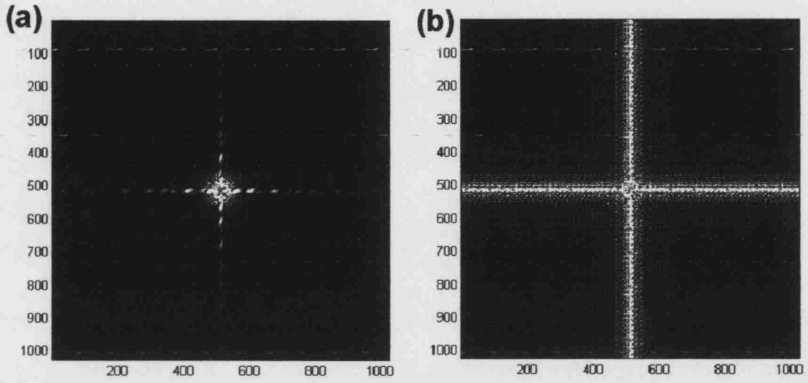

Figure 3.21: Simulated PSFs produced from the apertures in Figure 3.19(a) and (b) after using a coronagraph (solid disk and 80,20 Lyot stop). Images are scaled to same cut levels.

The coronagraph suppression performance could be significantly improved by using a continuous face-plate adaptive mirror compared to the segmented mirrors. Using a different Lyot stop, which masks the segment areas should help suppression performance, however such a mask requires extremely accurate alignment and would further decrease system throughput.

3.3 Summary

Producing a computer simulation that accurately represents a real-life system requires a level of program complexity and array size that is beyond the scope of current computing capabilities. In the real system there will nearly always be phenomena that have not been accounted for in the computer models and these must be taken into consideration when making performance predictions based on the simulated results. The computing limitations restrict the resolution of a model (array size) so that important effects may be under estimated, for example the DM segment gaps in the original osca.c program.

However, within the same model, the simulations do give reasonable results in comparative experiments such as the performance difference between the different sized masks (for both the focal and pupil masks) and the benefit of the different shaped masks. The simulations are also an invaluable tool to help determine the effect of a variety of external influences on the system – such as AO correction and atmospheric seeing conditions – and thus provides a basis on which to establish the optimum operating parameters.

CHAPTER 4

Testing and Commissioning

The design and building of a facility class instrument is subject to reviews and criteria testing before it is considered fully commissioned and available for general use. OSCA underwent two design reviews and criteria were drawn up which OSCA had to meet to be accepted at the telescope. These included more general standards such as system throughput, temperature stability and component lifetimes and other conditions more specific to a coronagraph, these will be outlined in the relevant discussions.

4.1 Installing the focal plane masks

A crucial alignment within the system is that of the focal plane (occulting) masks. The performance effects of decentring the occulting mask over the PSF were first discussed by Malbet (1996). This stated that the PSF-mask centring error should be less than 10% of the size of the mask to avoid significant performance degradation. We chose a 5% decentring error as our target with 10% as the limit of our acceptability. Our focal plane masks were deposited on wedged glass substrates and so an angular alignment error of $\pm 3°$ was also required to ensure the exit beam shift was within tolerances.

The apparatus used for the substrate alignment in the filter wheel is shown in Figure 4.1. Trial and error found the best method for accurately aligning the masks and keeping them in position. A fast-curing liquid UV-glue (Norland 81) was used to stick the substrates

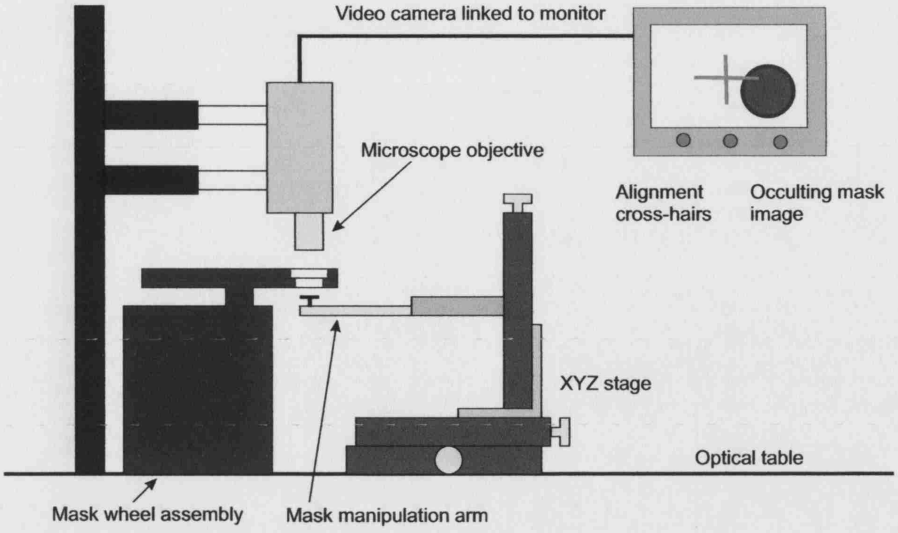

Figure 4.1: Diagram showing the basic set-up for aligning and installing the focal plane masks. A photograph of this procedure is shown in Figure 4.32

into the wheel. The best method for doing this was to place a few drops of the glue around the circular recess in the wheel using a syringe and then use a cotton-tip to spread it completely over the lip the substrates will sit on, leaving a very thin coating of the glue. Using a pair of optics tweezers a mask substrate is then carefully placed into the wheel recess, to be as centrally located as possible. Since the apex of the wedge angle had also been marked on the edge of the substrates, these were also orientated so that they all pointed roughly towards the centre of the wheel. For the first mask substrate put in the wheel the UV lamp was then turned on to cure the glue for a number of minutes. Although the glue sets after about 30 seconds, longer UV exposures allow the glue to fully cure for immediate and improved stability. Even and intense UV illumination is also important otherwise the glue may harden on one side of the substrate faster than the other, causing the positioning of the substrate to shift and lose alignment.

With all the apparatus firmly and heavily clamped to the optical table, including the mask wheel assembly, the detent arm mechanism on the wheel is pressurised to lock the wheel in place. A video camera with a microscope objective attached is focussed above the mask and aligned so that the occulting mask is central on the monitor screen. The centre of the mask is carefully measured and marked on the screen with cross-hairs. The wheel is unlocked and turned so that the mask is no longer under the video camera but positioned over a laser pointer instead. The beam passing through the mask produces two visible

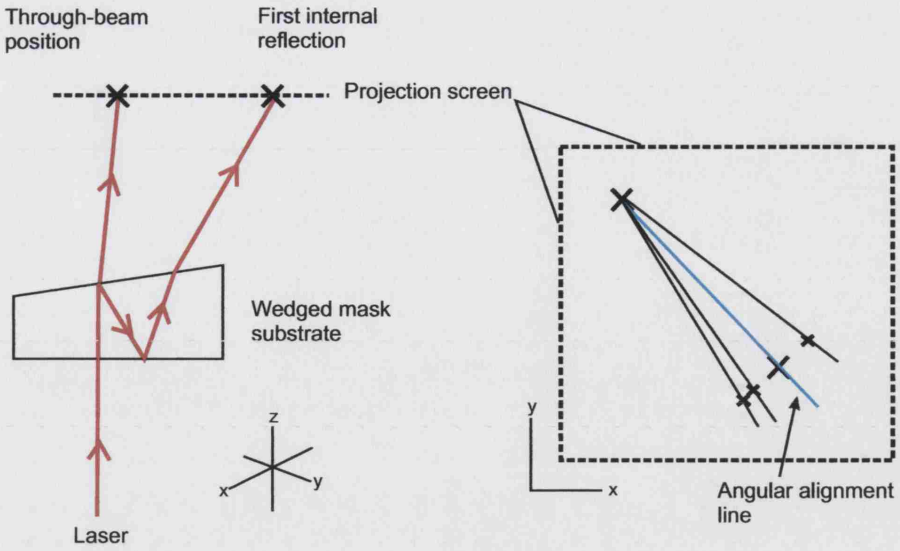

Figure 4.2: Diagram illustrating the angular alignment procedure for the wedged substrates in the filter wheel. The internal reflections follow the line of greatest angle in the substrate and so can be used to precisely align all the substrates. As the substrate rotates about z then the position of the first reflected beam on the projection screen rotates about the straight-through position as shown on the right.

spots on a projection screen fixed above the apparatus. The brightest spot corresponds to the straight-through beam whilst the second spot is much fainter and corresponds to the first internal reflection (see Figure 4.2). The position of the spots is marked and a line drawn through them, this provides the reference line for the angular alignment of all the wedges in the wheel. When new masks are installed in the wheel they are first orientated so that the angular alignment matches to the reference as close as possible, and then with the wheel locked in place with the detent arm and the mask viewed with the video camera, the substrate is manipulated with the aid of a custom built arm attached to an XYZ stage as shown in Figure 4.1. A flat, plastic screw attached to the arm sits under the substrate to lift it a fraction in the z-direction (so that it still remains in focus on the screen). Using adjustments in the x and y-directions the mask is moved so that it is positioned over the cross-hairs marked on the screen. This is an iterative process since placing the mask back into the recess (movement in z) can also move the mask slightly in x and y. Aligning the mask in this way does not tend to alter the angular position although this is double checked before curing. Repeated checks with the original reference are also made to ensure

that none of the pieces of apparatus have moved during mask installation. In this way the centres of the masks were aligned to within 15μm of each other (~1 pixel on INGRID).

Since the focal plane masks are deposited on wedged substrates, the angular alignment of the substrates in the wheel is important. This is because any change in angle would cause the position of the pupil plane to move causing an error in the alignment of the Lyot stop as shown in Figure 4.3. Since the Lyot stop is oversized to allow for alignment errors (particularly the vane masking) an angular error for positioning the substrates can be calculated.

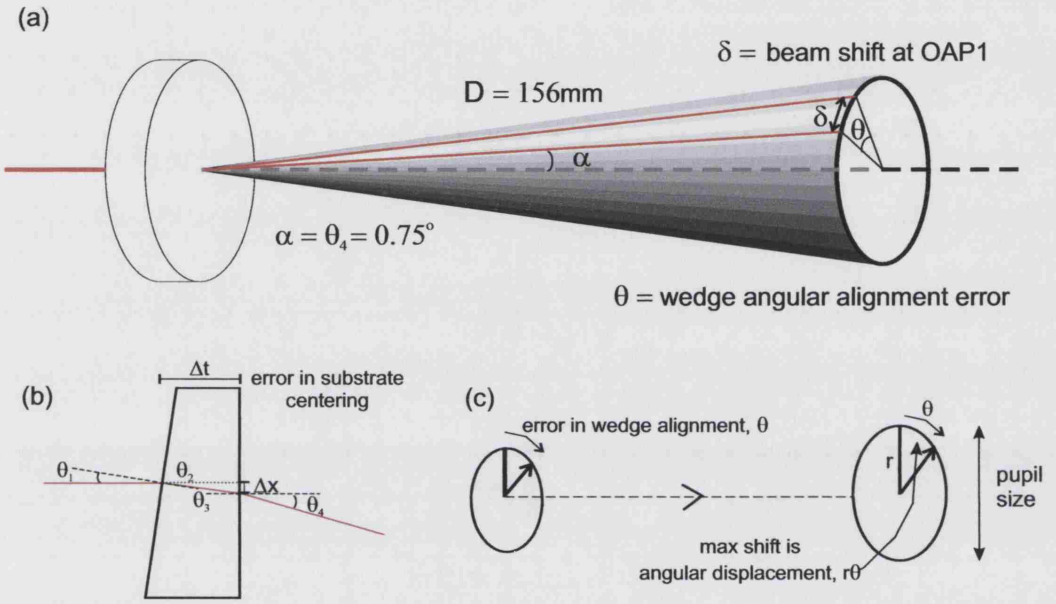

Figure 4.3: Diagram showing how (a) angular errors in wedge alignment and (b) centring errors in substrate alignment translate to position errors in the pupil plane, (c) shows how in addition to the shift δ in the pupil plane, the angular errors also correspond to a rotation of the pupil.

A perfectly aligned wedge should cause the beam to deflect in the x-direction (left-right) only. A misalignment in the substrate by a rotation of the wedge causes the beam to deflect by the same amount but in a different direction (x and y) and at the same time results in a rotation of the beam. To maintain vane masking (Lyot stop vanes are 0.37mm wide), a maximum allowable shift in the pupil position is 0.18mm; using equation 4.2 this corresponds to a wedge misalignment of 5°. A wedge misalignment by a given angle, θ, will additionally cause the pupil to rotate by $r\,\theta$. With reference to Figure 4.3(c), the outer

edge of the Lyot mask is positioned at $r = 3.9$mm, giving a displacement arc of 0.34mm for a $\theta = 5°$ misalignment. The result is that the telescope vanes becoming unmasked from the outer edge to about 80% along the length of the vane. If vane masking is not required only the error in primary masking is considered and the maximum shift in the pupil plane (for an 80% undersized mask, allowing $\pm 10\%$) is 0.5mm. This corresponds to a wedge misalignment of 14°. Since the primary masking is unaffected upon rotation, the rotational error in this case is not important. So, *for a Lyot stop with vane masking, the largest error resulting from a wedge angle misalignment is in the rotation of the pupil plane.* To maintain full vane masking the alignment must be within $\pm 2.6°$ or up to $\pm 5°$ for partial masking. An alignment error of $\pm 3.8°$ gives approximately 50% masking of the telescope vanes.

The position error of the occulting spots on the substrate (~1mm difference) will cause a slight shift in pupil position as well, since in the direction along the wedge a slightly longer or shorter path-length, Δt, through the substrate will be taken as depicted in Figure 4.3(b). However this error is very small compared to the other errors described and can safely be ignored as is shown below.

Taking the tolerance in the wedge angle ($\pm 0.05°$) into account and using Figure 4.3(b) as a reference, Snell's Law states

$$n_1 \sin\theta_1 = n_2 \sin\theta_2 \tag{4.1}$$

where n is the refractive index and θ is the angle of incidence with respect to the normal of the surface. For the wedged substrates, $\theta_1 = 1.5° \pm 0.05°$. From equation 4.1, with $n_1 = 1$ and $n_2 = 1.5$, $\theta_2 = 1.0° \pm 0.03°$. $\theta_3 = \theta_1 - \theta_2$ to give a value $\theta_3 = 0.5° \pm 0.06°$. The exit angle and the angular deflection between incoming and outgoing beam, $\theta_4 = 0.75° \pm 0.09°$. The change in thickness along the wedge is 0.31mm, so for a ~10% difference in mask position there is a difference in thickness of $\Delta t = 0.03$ mm. This translates to a shift in the exit beam of $\Delta x = 0.3\mu$m. Similarly by this method, the shift in exit beam due to an extra thickness of 2mm from bonding the gaussian HEBS glass to the wedges is of the order, $\Delta x_{gaussian} = 0.02$ mm. This is only about 10% of the maximum shift allowed for vane masking and so is well within tolerance; the shift due to occulting mask position is so small it can be ignored.

The largest translational shift in the pupil plane due to the wedge angle misalignment is

Table 4.1: Angular errors in positioning focal mask substrates. The maximum r is taken at edge of the 80% Lyot stop. There is an additional pupil shift due to the occulting mask centring - this is shown in Table 4.3.

Focal mask	$\Delta\theta°$ from 0.25 mask	Shift of pupil plane	
		δ (mm)	max $r\,\theta$ (mm)
0.25	0	0	0
0.40	-1.5	-0.053	-0.102
0.65	-0.25	-0.009	-0.017
0.80	2.5	0.089	0.170
1.00	1.0	0.036	0.068
1.60	-4.5	-0.160	-0.306
2.00	2.0	0.071	0.136
Gaus A	2.0	0.071	0.136
Gaus B	2.5	0.089	0.170
Target	3.0	0.107	0.204

the result of the *direction* of the exit beam changing. Rotating the substrate about its centre causes the normal to the wedged surface to rotate and so following the refracted ray through the substrate also causes the exit ray to rotate to trace out a cone of apex angle $2\theta_4$, as depicted in Figure 4.3(a), when rotated about a full 360°. The beam is diverging until it is reflected by OAP1 causing it to become a parallel beam, thus the maximum shift at the pupil is calculated from the focus→OAP1 distance. In this calculation the effect of the apparent focus of the beam shifting in x-y (the focus is on the front surface of the substrate to give an apparent shift $\sim 3\mu$m) is ignored. From Figure 4.3(a), it can be shown that,

$$\delta = 2\,D\sin(\alpha)\tan(\theta/2) \tag{4.2}$$

and since the angles here are small, this can be approximated to $\delta = D\sin\alpha\tan\theta$, where α is the angle of the exit beam with respect to the entry beam on the substrate, θ is the angular alignment error of the wedge angle and D is the focus→OAP1 distance.

The measured angular errors in the positioning of the substrates are listed in Table 4.1 and were taken at the time of installation. Although the 2 arcsec mask was used as the

reference during alignment, subsequent calculations showed the average error angle either side of the 0.25 arcsec mask to be at a minimum and so the pupil plane should be aligned with this focal mask. It can be seen that most of the masks are aligned to within the acceptable errors discussed previously with two exceptions - the 1.6 arcsec mask and the target mask. The target mask is never used for observing and so will not cause a problem. The misalignment of the 1.6 arcsec mask is problematic. If this mask is requested for observations then the best solution would be to remove the OSCA Lyot mask and use the pupil mask in INGRID. Some damage to the 1.6 arcsec mask at a later date (see §4.6) means that it could be best to avoid using this mask completely.

To remove substrates that have been bonded into the wheel a chlorinated solvent such as methylene chloride (paint stripper) should be dropped around the edge of the substrate using a syringe (to avoid coating the surface with the chemical) and then left for some time (up to several hours). This should dissolve the glue and the substrate can then be carefully removed using tweezers, wiped down and washed in a bath of acetone or alcohol.

4.1.1 Cleaning procedures

Before installation a cleaning procedure was established on rejected focal plane masks (i.e. masks with defects in the photolithography process) to determine what types of cleaning the delicate masks can safely withstand. Rubbing over the chromium mask with a cotton tip soaked in ethanol or acetone, even very gently, risks damaging the deposit. Large particles of dust or dirt are best blown off the substrates using a large hand-held rubber bulb with a narrow nozzle. This produces a strong jet of air which can be directed where needed. A can of spray-air fitted with a straw nozzle was tried but this was not as effective and risks liquid propellant spilling onto the mask and leaving a deposit (this was observed on one occasion during testing). The mask substrate should then be washed in a high grade propan-1-ol and a fibre-free lens tissue dipped in the same should be dragged gently across the surface several times. The surface should then be thoroughly rinsed in distilled water which is then allowed to run off the surface. All remaining drops should be blown to the edges with the air-bulb or a clean air supply. The overcoated silver mirror surfaces are also best cleaned in this way so as not to damage the sealed protected surface. For stubborn particles of dust on the mask substrates, a cotton-tip soaked in propanol may be used to carefully rub the dirt away as long as the tip does not touch the chromium mask.

Since the gaussian masks are written *inside* the glass a firmer cleaning procedure all over can be used.

4.2 Laboratory testing the focal plane masks

When time allowed between the focal mask production and shipping to the telescope, some testing was carried out on the masks in the laboratory at UCL. All these tests were performed in the visible (613nm red laser light) as no other wavelength lasers were available. The gaussian masks were produced after OSCA was shipped to the telescope, so direct comparison testing with the actual chromium masks on OSCA were not possible except for the 1.0 arcsec mask which was also produced after the first commissioning. Direct comparison between all the masks at the telescope is not currently possible since there is not a visible wavelength camera installed and the gaussian masks are only designed for visible wavelengths. This situation should change when OASIS is fully commissioned and the current alignment problems with OSCA are solved.

4.2.1 Opacity testing of the chromium masks

The opacity of the chromium occulting masks was tested using a 613nm laser, a set of ND-filters for comparison and a video camera/monitor. The measurements were more qualitative than quantitative and involved comparing the brightness of the laser spot through the occulting masks with the brightness of the laser spot through various ND filters until the spot brightnesses matched: the occulting mask was then assigned this ND value. The masks were assigned the following values:

0.25 arcsec → 4.5

0.40 arcsec → 4.5

0.65 arcsec → 3-3.5

0.80 arcsec → 3-3.5

1.00 arcsec → >5.0

1.60 arcsec → 3.0

2.00 arcsec → 4.5

There were not enough ND filters to create a combination greater than ND5 to allow a proper measurement for the 1.0 arcsec mask, however the optical density (OD) satisfied

our requirement. The different optical densities arise from the fact that the masks were made in batches and had chromium layers of slightly different thickness deposited on them. The optical densities in the infra-red appear to be lower than these values and have been estimated (incorrectly) by the telescope staff, so are not quoted here. The values could be taken as worst case indications, however the method used did not take into account the considerable contribution of scattered/diffracted light across the mask to the measurement of transmitted light. The actual OD values would therefore be higher (i.e. lower throughput). Ideally a near-IR laser is required to produce a narrow enough beam – the FWHM of the beam must be significantly smaller than the size of the mask – so that the contribution of light from edge scattering/diffraction is smaller than the amount of light actually transmitted through the mask. As the focal stop size decreases the contribution of light due to diffraction increases, so that for the smallest mask a measurement of the opacity becomes extremely unreliable.

4.2.2 Performance testing the Gaussian masks

A basic coronagraphic system was set up on the laboratory optical bench, as shown in Figure 4.4, to test and compare the suppression ability of the gaussian occulting masks with some spare unmounted chromium masks that were made for OSCA.

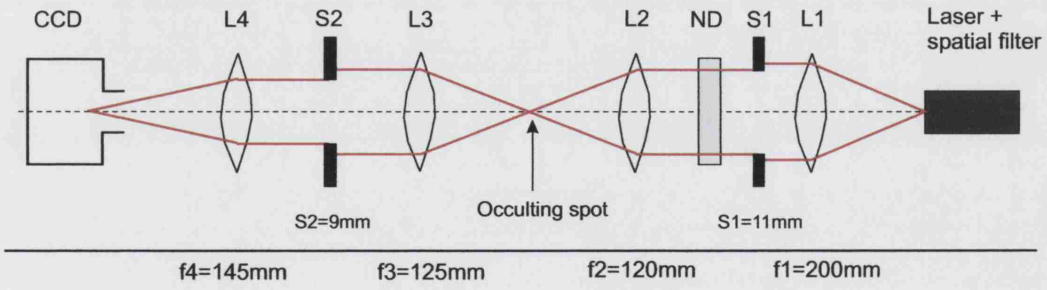

Figure 4.4: Optical experiment for suppression testing of coronagraphic masks. L indicates lenses and their focal lengths are noted underneath, S indicates a pupil stop (iris) with the aperture diameter noted underneath, ND indicates a neutral density filter. The CCD is an SBIG camera.

Two irises were used, one for the entrance pupil (11mm) and the second one 80% undersized (9mm) to act as the Lyot stop. Images were recorded at the focus for a variety of different occulting stops: both the gaussians, the 1 arcsec stop and the RAL template in lieu of the

other sizes (OSCA had already been sent to the telescope by the time the gaussians were manufactured). An SBIG camera was used to record the images. This is a 375 × 242 pixel (array dimension = 8.63mm × 6.53mm, pixel size = 23 × 27μm), Peltier cooled CCD. The camera was also moved to take pictures of the pupil in order to record the effect of placing different stops at the focus. All images were taken with automatic dark frames which were automatically subtracted in the SBIG software. Between each image a background image was also taken, i.e. the laser was turned off and an image of the same exposure time as the previous image taken. This was later subtracted from the data to account for any variations in background light levels in the laboratory. The data was analysed in MATLAB to produce radially averaged plots. The multi-rayed star pattern produced at the focus, particularly evident in Figure 4.5 taken without a Lyot stop, is the result of using a polygon shaped iris for the aperture stop. This figure and its corresponding graph (Figure 4.7) show the importance of a Lyot stop in suppressing the bright wings of a star which the occulting mask alone does not achieve.

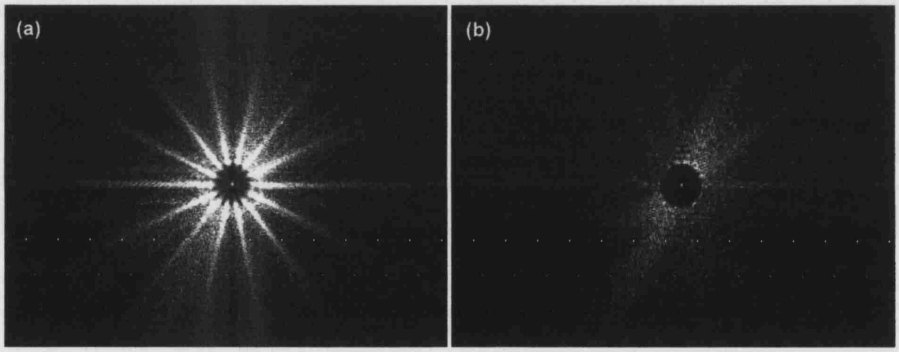

Figure 4.5: Difference in images using 1 arcsec mask, without (a) and with (b) a Lyot stop. Images are scaled to same levels.

From the graphs shown in Figures 4.7 - 4.10 it can also be seen that the 1 arcsec mask (later installed in OSCA) has an optical density equivalent to ND 5.5, in that the peak as seen through the mask is attenuated by a factor of 3×10^5 and this agrees with the opacity testing results. Suppression effects are lower for the other masks tested because the optical density is much lower, with a measured ND value from the graphs of about 2.3 for Gaus A, 2.0 for Gaus B and 2.3 for the 0.53 arcsec mask on the RAL plate. Additionally, these masks are about half the width and so remove a smaller proportion of the light at the focus. The gaussians should have the same ND value, which at this wavelength should be

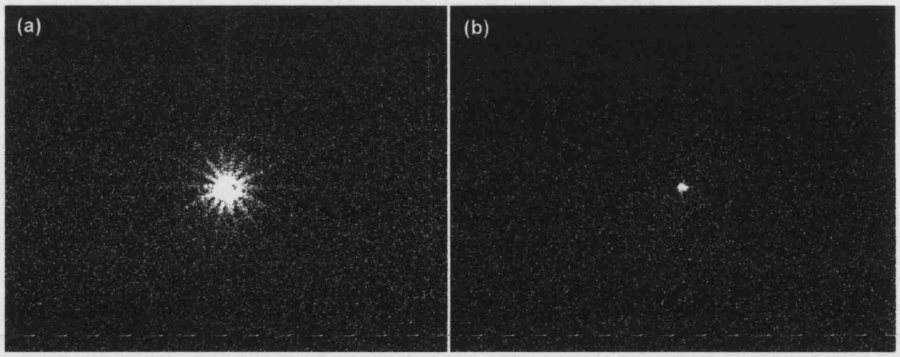

Figure 4.6: Difference in images using Gaus A mask, without (a) and with (b) a Lyot stop. Images are scaled to same levels.

3.0; the lower measured values could be due to a slight centring error. However, since Gaus A has a measured full-width half-maximum of 0.50 arcsec and a similar optical density as the 0.53 arcsec disk on the RAL plate, these two make a good set for comparing the effect of different shaped occulting masks.

Figure 4.8 shows the significant effect adding a Lyot stop has on the suppression with the Gaus A mask, the CCD count at 10 pixels from the centre is 5×10^3, compared to 1.2×10^4 for the 0.53 arcsec disk. Without a Lyot stop the value at 10 pixels from the centre for the Gaus A mask is 8×10^4 and for the 0.53 arcsec disk is 5.5×10^4. This shows that *for a gaussian mask at the focus compared to a solid disk of comparable size, an undersized pupil stop further along the beam works to greater advantage in suppressing the final image.* In terms of attenuation factors, the effect of adding the Lyot stop (measured at 10 pixels) with the Gaussian A mask compared to no Lyot stop brings about a factor of 16 drop in measured CCD counts, for the 0.53 arcsec disk this factor is only 4.6. At 60 pixels for Gaus A the factor is 1.83 and for the 0.53 arcsec disk 1.75, approximately the same. The extra effectiveness of the Lyot stop with the gaussian mask is therefore greatest close in to the mask, with the 'Lyot factor' dropping as the distance from the mask increases until it converges with the 0.53 mask values. This trend is shown in Figure 4.11.

A comparison of all the masks tested, plotting only the lines of mask + Lyot stop, is shown in Figure 4.12. From this it can be seen that Gaus A out-performs the 0.53 arcsec disk, particularly at distances close to the mask. Gaus B is worst, but this is likely due to a centring error on placing the mask over the focal spot. The 1.0 arcsec disk performs best

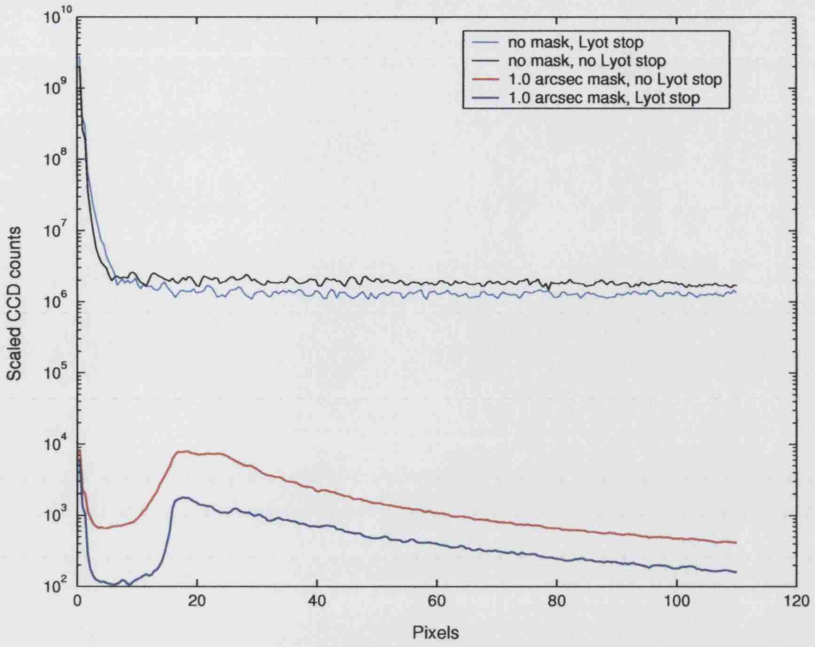

Figure 4.7: Radially averaged line profiles taken about the peak count, for different coronagraph configurations with the 1.0 arcsec focal mask.

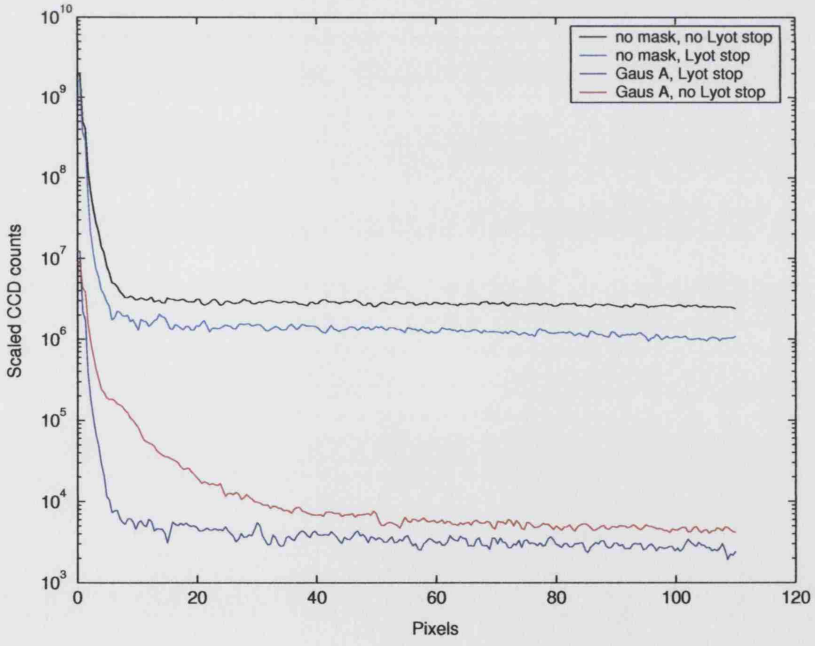

Figure 4.8: Radially averaged line profiles taken about the peak count, for different coronagraph configurations with the Gaus A focal mask.

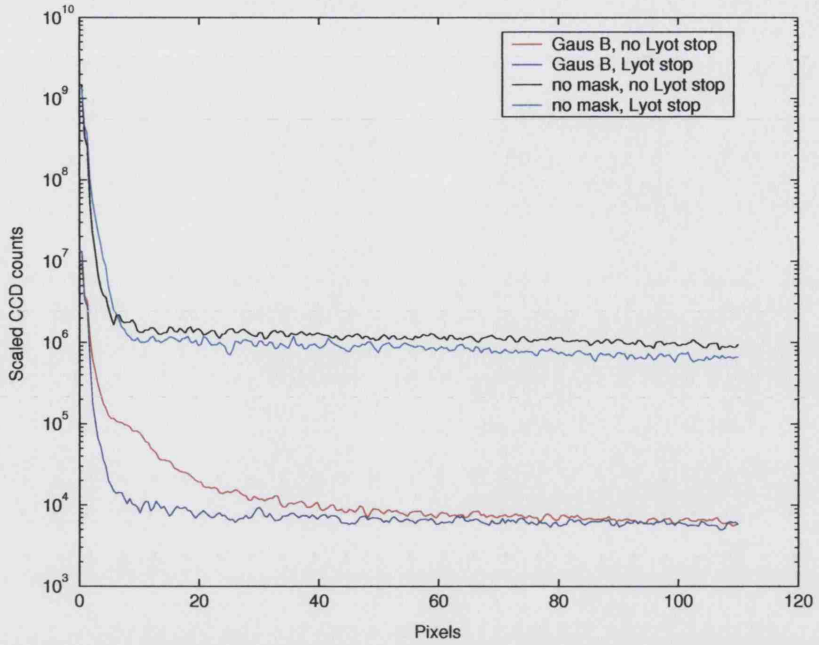

Figure 4.9: Radially averaged line profiles taken about the peak count, for different coronagraph configurations with the Gaus B focal mask.

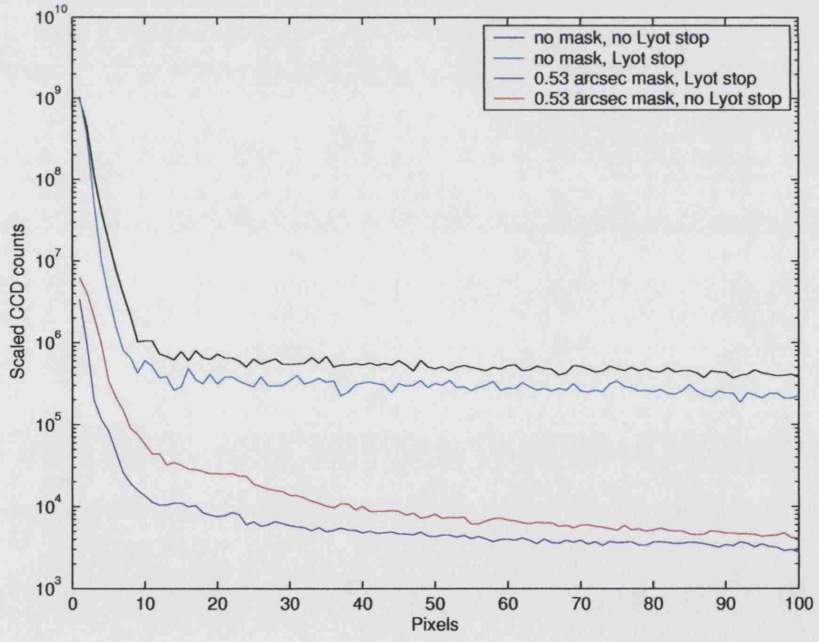

Figure 4.10: Radially averaged line profiles taken about the peak count, for different coronagraph configurations with the 0.53 arcsec focal mask.

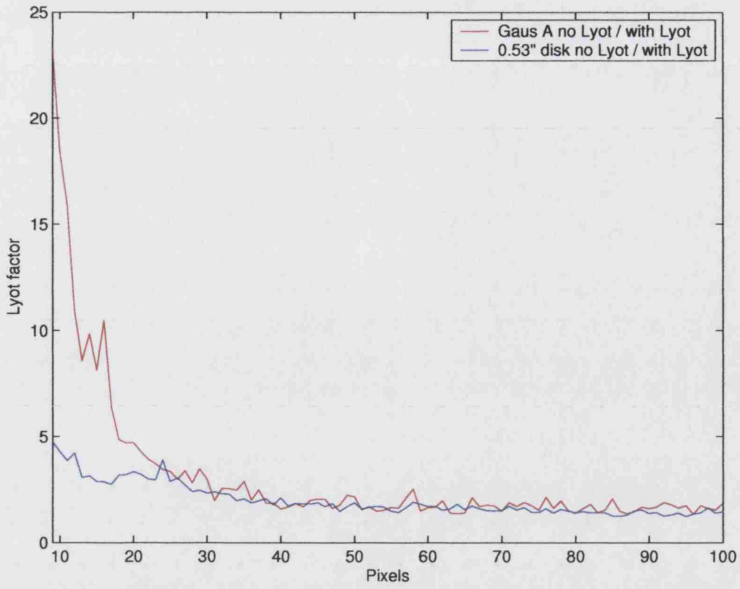

Figure 4.11: Plot of 'Lyot factors' for the Gaus A and 0.53 arcsec masks, where Lyot factor denotes ratio between the suppression curves with and without a Lyot stop. The pixel value starts a few pixels outside the outer edge of the masks.

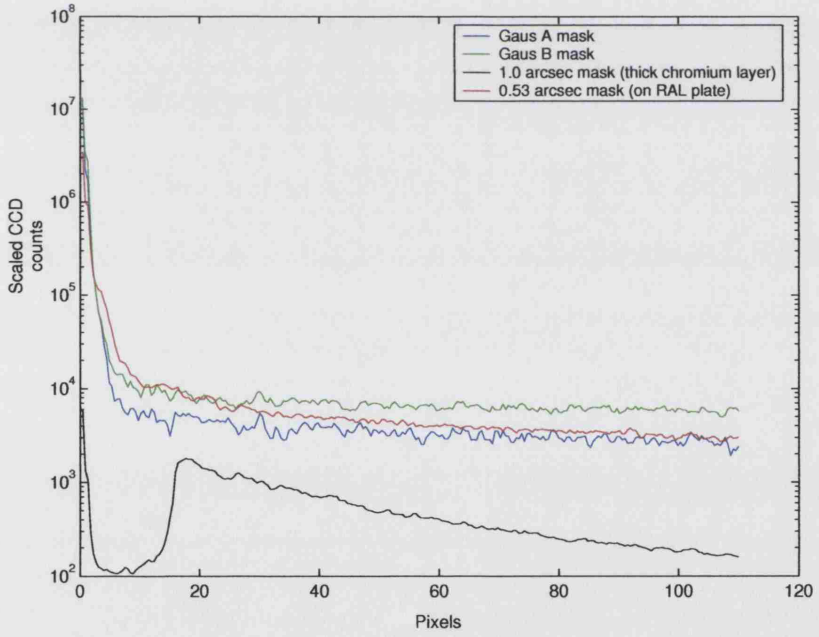

Figure 4.12: Graph comparing suppression of Gaussian masks with solid disks on the laser. All have the same Lyot stop and are scaled to the same exposure time and ND filter level.

of all, especially at larger distances from the mask edge. This is because it is twice the size of the other masks and also twice the optical density.

The pupil plane images were also analysed in MATLAB. To explain how the Lyot stop appears to have a greater effect when used with the gaussian mask, an image at the pupil plane was taken with the SBIG camera while using the Gaus A mask and the 0.53 arcsec disk. The Lyot stop would have a greater effect with the Gaus A mask if more light in the pupil were present in those areas masked by the Lyot stop. Since in this analysis it is the *distribution* of light that is of interest, the total CCD counts in each pupil image were scaled to 1 (normalised). The pupil images were then divided by each other and compared. Figure 4.13 shows the result of the normalised Gaus A pupil divided by the normalised 0.53 arcsec disk pupil. The result is very clear - there is a higher proportion of light around the edge of the pupil when using the gaussian mask. The brightest grey-levels in this picture correspond to a factor of 6.

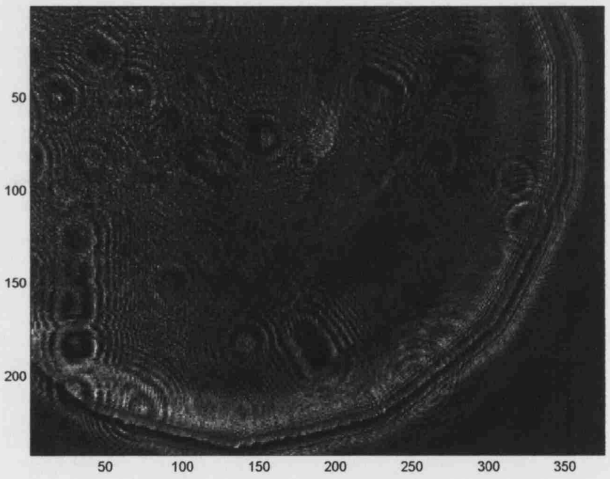

Figure 4.13: An image of factor differences between a normalised Gaus A pupil image and a 0.53 arcsec disk pupil, i.e. (gaus_pupil)/(disk_pupil). Laser speckle is present in the image.

To illustrate the theory described in the Introduction, the actual Lyot stops made for OSCA were put in place of S1 and images of this pupil plane were taken with and without an occulting spot at the focus. These images are shown in Figure 1.4. White light images of the gaussian focal masks (Figure 2.14) were also taken at this time and from these the

full-width half-maximum of their profiles were measured and checked against the design specifications.

4.3 Laboratory assembly of OSCA

All machined parts were assembled together with the commercially bought components and optics prior to anodising to check for errors. Anodising creates a very hard coating over the aluminium parts and so it is difficult to grind or machine components to correct for any errors discovered after this process. A photograph of OSCA in its 'bare metal' stage is shown in Figure A.1. At this initial assembly stage an infringement of the beam passing between OAP1 and OAP2 upon the mirror mounting block and clip for flat-mirror-3 was discovered. This meant that some of the mirror support and clip had to be cut away, so flat-mirror-3 is not held all around its edge as for the other two flats. Once anodised all the parts were re-assembled and tested for fit again, especially the smaller parts since anodising can alter the dimension by a small amount.

4.3.1 Alignment of opto-mechanics

Using the Zemax optical design as the initial guide, the optical components were adjusted, using a ruler and protractor for measurement, to roughly have the spacings in Table 4.2 and the input/output beam from OAP1 was checked to be at approximately 14 degrees. The orientation of the OAPs in their mirror holders were also checked, the mirror axis should be horizontal to the table and is marked by small pencil marks on either edge of the mirrors. The first focus falls on the *front* surface of the mask substrates. The base-plate of OSCA was then clamped firmly to the optical bench using the lines of holes in the table as an alignment guide. A laser beam was then set-up perpendicular to the OSCA base and the height of the beam set to match that of the NAOMI beam, 15cm above the optical bench. Since OAP2 does not have any height adjustment the height of the entire OSCA top plate is adjusted with the 3 ball-tipped locking screws (see Figure 2.19) so that the centre of this mirror is at 15cm: this corresponds to an 8cm top plate height and this should be checked to be the same at each end of OSCA to ensure that the top plate sits perfectly parallel to the bottom plate. Once this is completed OSCA can be locked into the 'up' position using the locking screw indicated for this purpose. Using the laser beam

as a guide, the remaining mirrors can be adjusted so that the beam falls on the centre of each of the mirrors in turn and exists the last mirror at a height of 15cm and along the same axis as the original beam.

Table 4.2: Spacing of optics in OSCA, in beam order.

From	To	Distance (mm)
Flat1	Focus1	104.3
Focus1	OAP1	155.7
OAP1	Pupil	133.5
OAP1	OAP2	309.0
OAP2	Flat2	69.1
Flat2	Flat3	21.5
Flat3	Focus2	63.1

OSCA is then dropped out of the beam by lowering the top plate. A spatial filter is attached to the laser and using a suitable lens and a 10mm iris an f/16 beam is created that approximates the NAOMI beam. A video camera is aligned on the other side of OSCA from the laser to record the focus of the f/16 beam. The camera is clamped to the bench as is the laser and the position of the focus marked on the monitor; the quality of the focus is also noted. The top plate of OSCA is locked up in place again and the Lyot mask is removed. With the OSCA base-plate unclamped, the whole assembly is moved up and down along the beam length until the beam is coincident with the centre of flat1 and OAP1 and the focus appears to fall in the vicinity of the wheel substrates. The base-plate is kept perpendicular to the beam. The 0.25 arcsec mask in the wheel is used to align the optics as this provides very little beam obstruction so that there is not a problem with visibility and it is also the mask whose alignment angle is closest to the average mask alignment angle.

When the focus is in the correct position a collimated, parallel beam is produced between OAP1 and OAP2. The collimation is checked by using a piece of flat glass plate a few millimetres thick mounted opposite a small white screen as shown in Figure 4.15. When placed in the beam the device produces a fringe pattern on the screen due to interference between the front- and back-face reflected beams from the glass plate. The beam is

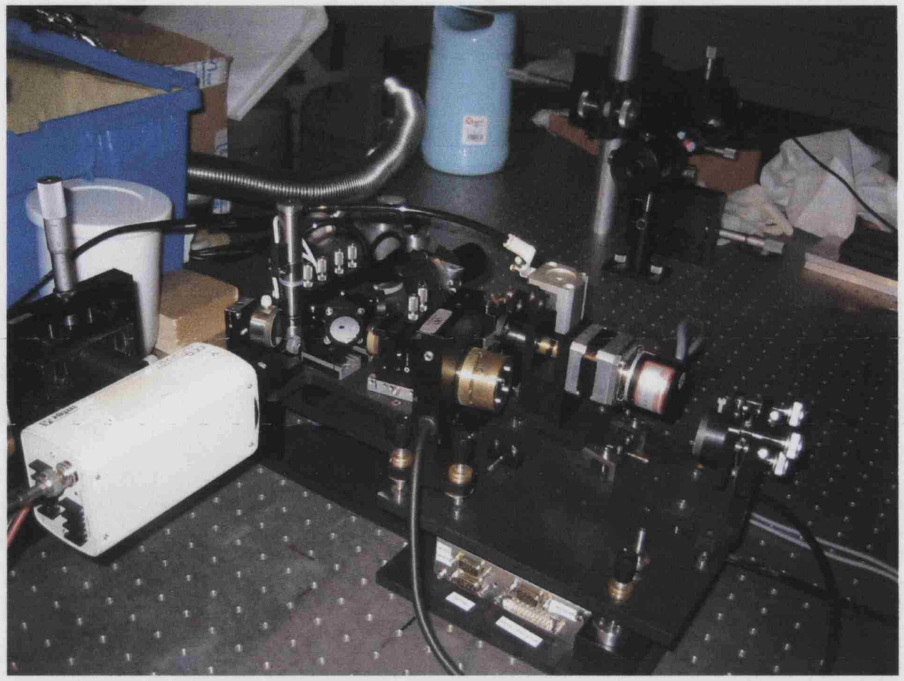

Figure 4.14: The internal alignment of the OSCA opto-mechanics. The laser and video camera can be seen either side of OSCA. Most of the adjustment screws discussed are clearly seen in this picture and the cut-away mirror clip for flat3 can also be seen.

collimated when the pattern disappears or appears as a random distortion due to the imperfect optical surfaces. This phenomenon occurs because both reflected beams retain coherence and add together to produce a maximum (bright), minimum (dark), or some level in between across the entire combined wavefront. A slightly converging or diverging beam will encounter a linearly varying path length across the wavefront coming from the back reflection through the glass which causes a corresponding linear phase change across the wavefront. This adds together with the front reflection and forms a fringe pattern.

The focus is correctly placed by adjusting the bench-mounted screw positioning blocks which allow the whole of OSCA to be moved up and down the beam in small increments until a collimated beam is produced. OSCA is then clamped down to the bench. The collimated beam from OAP1 should fall at the centre of OAP2. If it is not, the tip/tilt of OAP1 can be adjusted to place the beam correctly and the collimation then checked again since the adjustments may have altered the position of OAP1 with respect to the focus. The beam from OAP2 should fall just inside the right edge (as seen from the beam

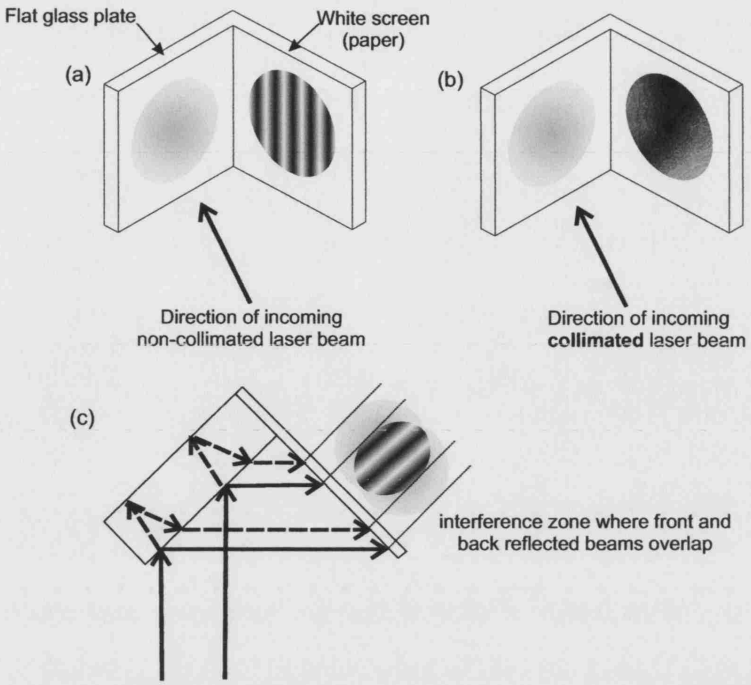

Figure 4.15: Diagram illustrating the construction of a parallel beam checker used for aligning OSCA. (a) and (b) show the type of pattern that is likely to be seen for a non-collimated and collimated beam, (c) is a cross-section from above showing the different beam paths.

direction) of the flat2 mirror although remaining centred vertically. The exit beam can be steered to the correct place on the video camera by adjusting the last mirror (flat3) mount; the beam should be along the same axis at the input beam. The spot is unlikely to be in focus at this stage. The focus can be adjusted by altering the path length between OAP2 and the required focal position. This is achieved by moving flat2 and flat3 forwards and backwards on their adjustable slide mounts. In doing this the angle of the mirrors also have to be adjusted to keep the beam in the correct position, so this is an iterative process. Each time a change is made the beam should be moved to the correct position on the video monitor and the focus quality checked until the best focus is achieved. At this stage the focus may appear slightly aberrated. This can be improved by making very small adjustments to the OAPs in an attempt to obtain a better matching of positions, although care must be taken to retain the collimated beam. Finally, once a satisfactory focus has been achieved, the mask-wheel plate can be adjusted using the pairs of opposing screws that move it across the top plate so that the occulting masks are moved in to focus

on the video screen. The reflected beams from the back of the mask substrates should be checked to ensure that they are not re-entering the optical system and if so a slight angular twist should be put on the wheel plate so that the reflected beam exits in a safe direction. The wheel plate is then clamped securely to the top plate. Finally, the Lyot mask is attached to the SR50 and with the focus behind the 0.25 arcsec mask the Lyot stop assembly can be adjusted using the 3 ball-tip locking screws until it is approximately centred in the beam. This completes the internal alignment procedure for OSCA.

4.3.2 Electronic systems integration

OSCA was taken to Durham University for a week at the end of January 2002 for integration and testing with the electronic and software control systems that had been built there. A full mechanisms check was performed, the base-level engineering control software tested and calibrated and a limited durability test was carried out on the stepper motor controlled devices. Due to limited time the durability testing was to leave the Lyot stop rotating over a full day and the mask wheel going to random mask positions for about half a day; neither failed during this time.

The mechanism testing checklist was as follows:

- Lyot stop: rotation, zero point, positioning, calibration.
- Mask wheel: rotation, zero point, encoder, calibration.
- Deployment stage: in and out switches (DC80).
- Detent arm: in and out switches.

In general the electronic systems integration was successful and no major problems were encountered. One thing was noted however, and this is that the Lyot stop has no way of telling if it has failed: the failure will only be discovered the next day when it is re-started since the software runs a full check on boot-up. However if a failure of the Lyot stop was suspected during observing (excess image artefacts from vanes noticed) then an image could be taken in INGRID pupil viewing mode to check the alignment.

The pneumatic systems were also well tested and the air pressure measured to find the minimum required to firmly lock the detent arm in place and deploy the top plate of OSCA.

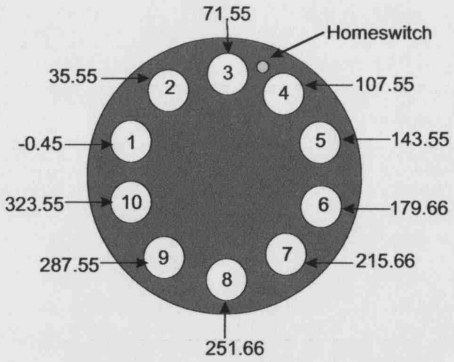

Figure 4.16: Mask positions on filter wheel as viewed from motor with the encoder positions indicated, as calibrated in January 2002.

4.4 Commissioning at the WHT 19 May 2002

OSCA and its electronics were shipped out to La Palma at the beginning of May 2002, with its first commissioning scheduled for a few weeks later. For future reference the dimensions of the shipping crate were $118 \times 51 \times 75$ cm and it weighed approximately 65 kgs, the electronics box being the largest component of dimension $40 \times 43 \times 51.5$ cm. The instrument arrived intact and the first task on unpacking was to thoroughly clean all the optical surfaces and reassemble OSCA from its main parts.

4.4.1 OSCA set-up at the test-station

An f/16 laser beam was set up on the optical test bench at the WHT using a He-Ne laser, spatial filter (pinhole + lens) and iris (aperture stop). A video camera was then aligned exactly opposite the laser so that the focus gave the smallest and sharpest image on the video monitor screen. OSCA was placed to interrupt the beam as it would in GHRIL. In this way the internal optical alignment of OSCA could be carried out, as described in §4.3.1, to minimise the adjustments that would need to be done in GHRIL where space and easy access to the different parts is very limited (see Figure A.6 for a picture showing the access restriction).

Once the internal alignment was complete the electronic systems were set up and plugged into OSCA. Care must be taken to plug all the connectors into the correct socket on the crate. The focal mask wheel encoder was accidentally blown in this way but has since

Figure 4.17: TV monitors relaying an image of the OSCA focus at the test-bench of the WHT. On the left is the focus with the 2 arcsec occulting mask in place and the Lyot stop. The picture on the right is the same set-up but with the Lyot stop removed. The effectiveness of the Lyot stop in removing the bright diffraction pattern is clearly demonstrated.

been replaced. The control software was then brought online and a full systems check performed. All the mechanisms were tested for correct response to the control signals and the detent arm, deployment stage, mask wheel positions, Lyot stop rotation rate and direction status signals were tested. The new software for OSCA then had to be integrated with the telescope EPICS records and checked to be compatible with the system. The OSCA control systems were then checked again but this time from the telescope control system.

4.4.2 Installation of OSCA at GHRIL

OSCA was carried up from the test lab and put in place on the GHRIL optical table between INGRID and the NAOMI wavefront sensor. There was a slight problem with the 3 adjustment screws on the back of the OAP1 mirror mount coming up against the wavefront sensor so that OSCA could not be moved into the correct position. As a temporary solution that panel of the wavefront sensor cover was removed allowing OSCA to move further back. Since then a small part of the cover has been cut away to make room for OSCA.

The NAOMI alignment laser was set up to check the height and path of the beam through

the OSCA optics. This is a pencil beam laser aligned along the NAOMI optical axis and it allows the height of the OSCA optics to be checked and adjusted if necessary so that the beam falls upon the centre of them. It also provides a rough positioning aid for the placing of the entire system on the bench so that flat1 picks up the beam and directs it to the correct position on OAP1.

The next stage is to remove the laser and turn on the NAOMI artificial star (white light behind a pinhole) to establish the focus of the system. The NAOMI simplexing procedure must be performed in order to obtain a reasonably flat mirror and an acceptable focus. Since the artificial star is a white light source the collimation between OAP1 and OAP2 cannot be checked. It is important therefore that flat mirror1 and OAP1 are not adjusted otherwise this collimation is likely to be lost. Since it is the *position* of the focus leading to OAP1 that defines the collimation, the occulting masks must be used as the reference. Its is therefore essential that the position of the focal masks along the beam is not altered either. With OSCA out of the beam, the NAOMI focal point is brought into focus on INGRID and the position and quality of the focus recorded (an image taken). OSCA is then deployed into the beam and the largest (2 arcsec) mask locked in place – this is the easiest to see in an out of focus image. OSCA is then moved up and down the beam until the best (probably out of focus) image of the mask is made on INGRID. The path length through OSCA from OAP2 can be lengthened or shortened by adjusting flat mirrors 2 and 3 on their slides which is done to improve the focus of the occulting mask on INGRID. This does not mean that the first focus is placed at the occulting masks yet, and so OSCA should be moved slightly up and down the beam again to improve the focus of the mask on INGRID again. This is an iterative process, and each time a reasonable focus is obtained with an occulting mask, a smaller mask should be selected and the focus checked and adjusted as described above until the target mask can be seen in sharp focus. At this point the NAOMI artificial star should also be in focus. The exit focus is checked by taking an image with INGRID and rotating the OSCA base plate as necessary to bring the focus to the correct position on the detector (the pixel position of the focus without OSCA should be marked to provide the correct location).

The base plate of OSCA is then clamped down to the GHRIL bench. Since the INGRID detector is built from 4 quadrants, it was desirable to place the OSCA occulting mask position away from these boundaries. This was achieved by rotating the wheel assembly

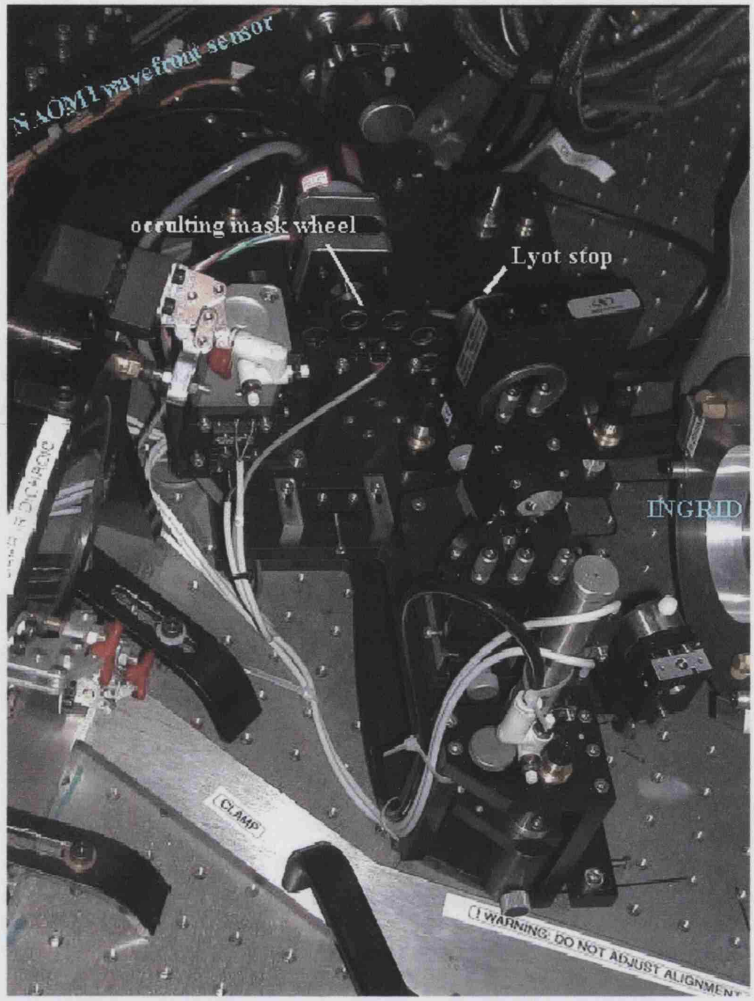

Figure 4.18: Photograph of OSCA after initial alignment on the GHRIL bench at the WHT. The cover of the NAOMI wavefront sensor had to be removed due to an impingement with the OAP1 mirror holder.

which is mounted over two ball-bearings with a ball-tipped screw opposite to provide tilt, and locking it in place by 3 screws to the wheel-plate. The mask wheel should not be moved so far that the beam is vignetted.

The next step is to check the OSCA pupil and align the Lyot stop to the telescope pupil. The 0.25 arcsec mask must be locked in the wheel position for this alignment to minimise pupil shifting errors. The (artificial) star should also be positioned behind the mask as this defines the beam axis through the system. INGRID is switched to pupil viewing mode and the telescope primary mirror cover opened - this was found to be necessary to obtain the correct lighting to see the pupil. On viewing the pupil with OSCA out of the beam it was

discovered that the telescope, NAOMI and INGRID were not aligned with each other and were mismatched by almost 30%. OSCA requires that *all* pupils along the entire optical path are aligned with each other to 5% or better. The performance of OSCA depends on the pupil plane masking as well as the focal masks and so it is essential that this pupil alignment is carried out. The ING and NAOMI team who were also present at the OSCA commissioning performed this alignment which is a fairly lengthy procedure and took nearly 2 days. After this was completed the OSCA Lyot stop alignment procedure could be started.

With the Lyot mask removed and INGRID in pupil viewing mode, the OSCA pupil should be checked to coincide with the INGRID pupil and there should be no brightness gradient over the pupil to suggest that the OSCA beam is entering at an angle. Flat3 can be adjusted if needed until the pupil is aligned with INGRID. If the position of the focus needs to be moved then this can be achieved by moving flat3 on its slide adjuster, however flat2 will also need to be moved on its slide adjuster to re-focus.

The telescope secondary is centred over the central 4 NAOMI segments, and using the (80,20) Lyot stop the corners of these 4 segments can just been seen around the central obscuration when it is properly centred. With the Lyot mask mounted on the SR50 rotation stage, the adjustments to centre the stop are made using the 3 ball-tipped locking screws. These screws seat the SR50 plate to kinematic pads attached to the OSCA top-plate. The adjustment is very awkward to make since there is little space to manoeuvre around the screws and care must be taken not to touch the mask-wheel that is very close to one of the adjustment screws. Only small changes in Lyot mask position can be made because the screws only provide a rotational displacement or height adjustment. However if the alignment has been correct up to this point a small displacement is all that is necessary. The SR50 plate is locked in place with some additional long M3 screws on each side of the stage.

To test the Lyot stop alignment, the Lyot stop was rotated by 360 degrees and an image in pupil viewing mode was taken every 90 degrees. The pixel position of the centre of the Lyot stop was measured in each image and is plotted in Figure 4.19. The centre of this circle is the true centre of rotation of the SR50 and for a perfect alignment the Lyot mask would need to be shifted slightly on its mounting until the centres matched. The centres are misaligned by approximately 2 pixels, corresponding to 0.12mm at the OSCA pupil.

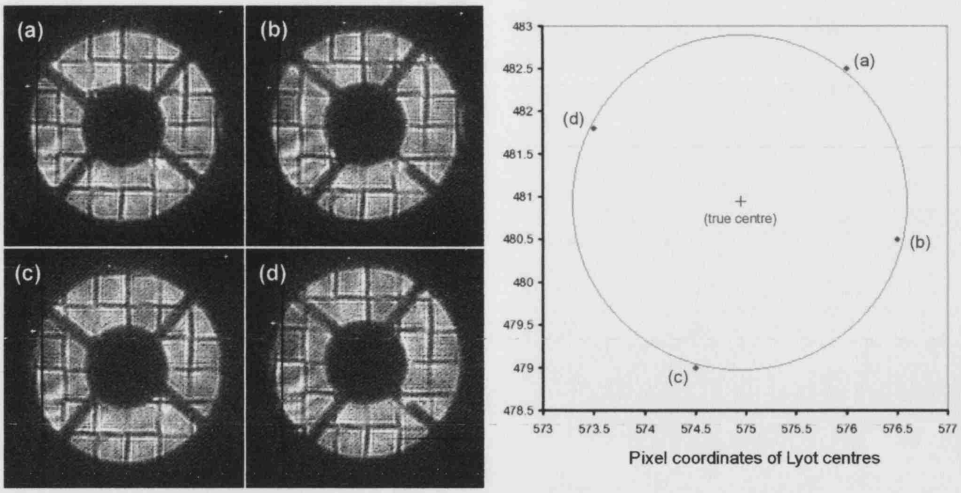

Figure 4.19: De-centring error of Lyot stop mounted on SR50 rotation stage. The red cross marks the true centre of rotation to which the mask should be aligned.

This error is 30% less than the maximum acceptable shift to maintain vane masking and so no further adjustments were initially made due to a shortage of time. However these were performed at a later date. Initially the INGRID pupil viewing mode was not available in the GHRIL optics room and alignments had to be directed via phone from another person viewing the image in the Control Room which made the procedure very arduous. When this was later remedied by the addition of a computer in the GHRIL room the procedure was much speedier.

The Lyot stop must also be aligned to match the telescope vanes and to ensure that the Lyot stop rotation is tracking correctly with the telescope derotator from the information provided by the TCS (telescope control system). Due to the alt-az mounting of the WHT, the image field at any of the foci will rotate as the telescope tracks it across the sky. To keep the image stationary at the Nasmyth platforms a set of mirrors mounted on a motor driven rotation stage are used, so that the image rotates at twice the angular speed of the rotator optics. A side effect of the derotator is that the pupil image of the telescope rotates so to maintain the position of a Lyot stop with vane masking the Lyot stop must also rotate at the same rate as the pupil. The Lyot angle is given in Equation 4.3 and is counted from the built-in home switch in the SR50, where the smallest step size is equivalent to an angle of 0.006 degrees.

$$\text{Lyot angle} = 2 \times (\text{Derotator angle}) - \text{Zenith angle} - \text{offset} \tag{4.3}$$

The derotator and zenith angle are passed to the Lyot stop control software via the TCS. The offset angle is the alignment angle needed to initially align the Lyot mask to the telescope vanes and is not changed unless the Lyot mask is manually moved (e.g. changed for a different mask).

4.4.3 Testing OSCA at GHRIL

For full acceptance at the WHT, OSCA was required to undergo testing to ensure it had met its performance criteria as was specified in the Critical Design Review.

Repeatability

For the OSCA deployment system the requirement was that the stage should locate to the original position to within 50μm every time. The stage was deployed in and out of the beam 8 times and the position of a corner on the target alignment mask measured. The position of the mask was found to relocate to the same pixel position each time. Since the pixel resolution corresponds to 13μm per pixel this is well within the specified limits.

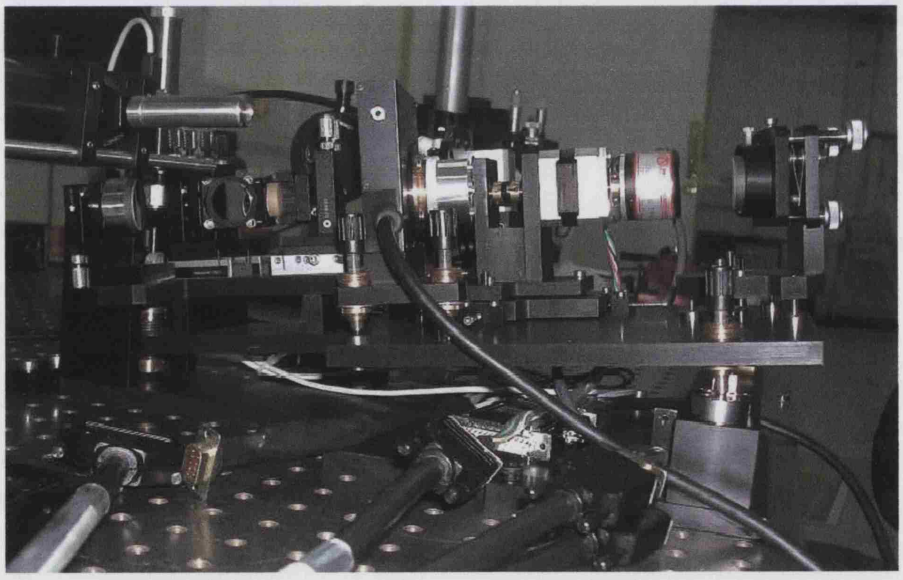

Figure 4.20: The OSCA deployment mechanism. Picture shows OSCA in the 'out' position.

To acquire a new star when observing with NAOMI the system is set to open-loop (the DM

stops receiving correction signals from the wavefront sensor), the telescope slews to the new object and is then guided interactively until the new target appears on the wavefront sensor and NAOMI closes the loop again. The new star will not be in the same position on INGRID as the previous target, so it must be re-centred behind the OSCA occulting masks for each new target. It is the adjustments to the NAOMI pick-off probe that allow the star to be centred accurately, so these x and y offsets are recorded and used again if the target needs to be reacquired. It is then more likely that the star will return to the correct position with no re-centring on the mask required.

To reduce the time needed for re-alignment after a focal mask change the OSCA occulting masks were centred to within 10μm of each other. This allows an observer to try out a different mask on a star with minimum on-sky time lost.

Each occulting mask was selected in turn and the position of the centre of the mask measured. The results of this test are listed in Table 4.3.

Table 4.3: Performance test results - occulting mask repeatability. The pupil shift due to the centring differences have also been included, the 0.25" mask is the reference.

Mask size (arcsec)	Pixel position	Pupil shift (μm)
0.25	637,536	0
0.40	637,536	0
0.65	636,536	14
0.80	637,536	0
1.60	636,537	20
2.00	636,536	14
Target	637,537	14

This gives the average position to be (636.6,536.3), the maximum deviation from this is for the 1.6 arcsec mask and corresponds to a distance of 0.92 pixels or 11.9μm. The maximum shift between masks is 1.4 pixels or 18.1μm which is outside the 10μm limit. Due to the critical nature of star centring behind the occulting mask and the nature of acquision with NAOMI, which makes it likely that in most cases the star will need to be re-centred anyway, the mask positions were deemed to be acceptable. The maximum shift in the

pupil that would result from the beam having to be in different positions for the different occulting masks is only 5% of the Lyot vane masking size which is also acceptable.

Throughput

System throughput tests were performed in J and H bands with and without OSCA in the NAOMI beam. Accounting for changes in throughput due to the Lyot stop (53%), this gives the percentage *loss* of throughput compared to the straight through NAOMI beam (losses from scattering due to dust on surfaces, reflectivity of mirrors and transmission of windows in OSCA) as follows:
J band - 11% ± 2%
H band - 9% ± 5%

These results are better than expected given the manufacturing data. Considering the 5 mirror surfaces ($0.97^5 = 0.86$) and the wedged window ($0.86 \times 0.96 = 0.82$), the expected value is around 82% throughput or 18% losses for a beam passing through OSCA.

Measurements in Ksh have not yet been made.

Suppression

Off-sky suppression tests using the artificial star were not possible at the time of commissioning. This is because the pinhole was fixed in position and so could not be positioned behind the OSCA occulting masks. This did however become possible the year after commissioning (see §4.6, Figure 4.38) when the pre-NAOMI filter change mechanism became available for use.

4.4.4 On-sky testing

Only two nights on-sky were allocated for the OSCA commissioning so time was very limited. The night of 23rd May was plagued by extremely bad seeing throughout (5 arcsec recorded at worst) and the AO system could not be used. Since the largest occulting mask in OSCA is 2 arcsec it was also not feasible to do any performance testing.

The last night (24th May) saw very variable seeing over the course of the night and sky-

location and the presence of high cirrus cloud also caused problems on occasion. NAOMI was used although centring objects on the OSCA occulting masks was difficult when the seeing was bad and OSCA performance was degraded. The average seeing was 1.5 arcsec, and with the AO system an average corrected PSF width of 0.5 arcsec was obtained.

Table 4.4: Radially averaged suppression factors from initial commissioning on-sky, measured at 0.5 arcsec from outer edge of mask.

Mask size (arcsec)	Corrected PSF size (FWHM, arcsec)	Wavelength band	Suppression (without/with OSCA)
2.00	0.40	J	2.0
2.00	0.45	Hcont	2.0
0.80	0.60	H	1.5
0.65	0.60	Ksh	1.6

Measurements of the suppression were obtained in J, H and Ksh for some of the masks and are shown in Table 4.4. Pixel counts were measured using a radial averaging program with values taken at approximately 0.5 arcsec from the edge of the occulting mask. It should be noted that in directions away from the bright diffraction spikes (caused by the NAOMI square segments) the suppression factor is better than the averages recorded in Table 4.4. Figure 4.21 shows a cross section taken in such a favourable direction. In these directions OSCA obtained a suppression factor of around 3.5.

Due to the poor seeing, masks smaller than the 0.65 arcsec mask were not tested. Generally it is not advisable to use masks less than half the size of the uncorrected seeing since there will still be an AO halo of this size about the corrected image and the performance of the coronagraph is dependent on the amount of light blocked from the central star. Essentially, there must be a reasonable Strehl ratio for the corrected PSF to give good coronagraph performance. The importance of this is highlighted in §4.6.

Since OSCA is not a cooled system the occulting masks can be seen to glow in K band. This can be seen in Figure 4.22, a sky-image with the 2.0 arcsec occulting mask. More tests need to be undertaken to determine to what extent this affects the performance of OSCA in K.

A pupil image was taken with the 1.6 arcsec occulting mask at the focus (Figure 4.23).

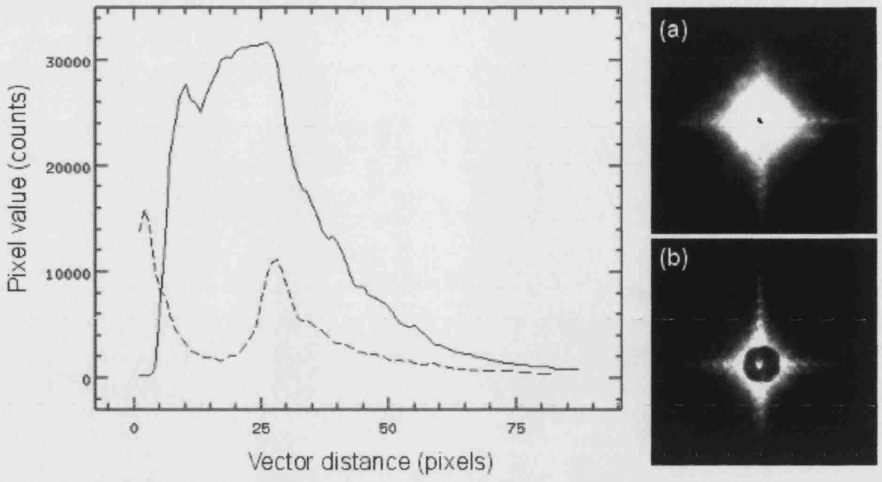

Figure 4.21: Cross-sections taken in the same high-contrast direction from H-band images (a) without OSCA (solid line) and (b) with 2 arcsec mask (dashed line). The peak at zero is due to the OSCA mask having a small throughput. The sharp drop at zero with no coronagraph is due to detector saturation. The INGRID pixel scale is 0.04 arcsec/pixel.

This is the occulting mask calculated to have an unacceptable error in its angular alignment based on the shift it would cause in the pupil. Measurement of the coordinates of the centre of the Lyot mask and the centre of the pupil plane (centre of NAOMI segmented mirror) show that the Lyot mask is off-centre by 0.46mm. Taking into account the possible misalignment of the Lyot mask over the pupil by 2 pixels (120μm) this gives an approximate shift of 0.34mm. The estimated minimum error is 0.28mm (160μm from Table 4.1 + 120μm) based on the Lyot alignment with the 0.25 arcsec mask in place. It is possible that another mask was in place during the alignment in which case the error could be larger.

4.4.5 On-sky target observations

Attempts were made to observe science targets during the course of the night. A list of suitable and interesting astronomy targets with possible calibration/subtraction stars for each was drawn up in the weeks prior to the commissioning. Tables 4.5 and 4.6 show the list of 15 objects that were compiled, however due to limited time and poor seeing

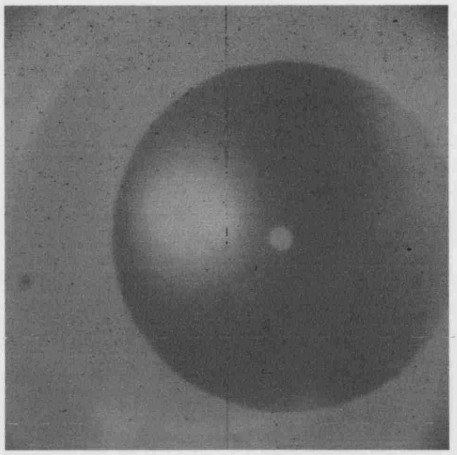

Figure 4.22: Sky-image in K-band showing the thermal emission of the non-cooled 2.0 arcsec occulting mask. The additional bright circular glow to the left of the mask is not from OSCA, but thought to originate from the INGRID fore-optics or the de-rotator.

conditions only the targets in Table 4.5 were observed. The objects were chosen based on their need for coronagraphic observations, i.e. features that would not otherwise be easily observable and would demonstrate the benefits of using a coronagraph. They also had to have a V-band magnitude of less than 12 (i.e. brighter) due to the sensitivity of the NAOMI wavefront sensor and be observable during the night at an altitude (z) greater than 40°; below this the turbulence is generally higher due to the high air-mass and ideally AO observations should be done for $z > 50°$. Additionally two possible subtraction stars were found for each target. Subtracting an occulted calibration star from the science target provides an additional increase in sensitivity. A full description of the OSCA data reduction technique is given in §4.7.1. These subtraction/calibration stars were selected to be as close to the target object in all respects – sky position, colour/spectral type and V magnitude. The attempt was to try and select single stars (no known companions) with a 10 arcsec field about them that is free of any other (particularly bright) stars. Finding good subtraction stars is difficult as there are no comprehensive catalogues for this and information can be incomplete or incorrect in the ones that are available. The subtraction stars chosen here were selected with aid of the SIMBAD astronomical catalogue.

The type of objects on which observations were attempted fall under the following categories:

Table 4.5: Observing targets for May 2002 commissioning run

Ref #	Object	RA/DEC	Calib star RA/DEC	V mag	Obs time	Obs bands
1	WR 124	19 11 30.88 +16 51 38.2	19 12 18.01 +16 50 54.3 19 09 46.07 +17 13 35.5	11.08	02:30 ...	J & H broad
2	P Cyg	20 17 47.2 +38 01 58.5	20 17 47.20 +36 48 22.7 20 09 25.62 +36 50 22.6	4.8	03:00 ...	FeII, H-cont, Pa-beta Br-gamma, K-cont, J-cont
3	HD 168625	18 21 19.55 -16 22 26.1	18 21 14.89 -16 22 31.8 18 22 50.10 -16 55 31.7	8.44	03:30 ...	J & H broad
5	GJ 623	16 24 09.3242 +48 21 10.457	16 26 21.32 +48 57 53.7 16 21 47.22 +47 43 49.6	10.27	23:00 ...	H & K broad
7	HD 141569	15 49 57.75 -03 55 16.36	15 57 08.63 -06 17 47.9 16 05 08.49 -03 31 39.4	7	00:00 ...	J & H broad
9	Y Cvn	12 45 07.83 +45 26 24.92	12 54 56.52 +47 11 48.2 13 05 52.27 +45 16 06.8	5.3	22:30 - 02:30	J & H broad
15	HD 150451AB	16 41 11.5441 -01 00 01.122	16 41 16.73 +01 14 43.6 16 50 22.25 -02 39 15.7	6.3	00:00 ...	J & H broad

Table 4.6: Other observing targets compiled for May 2002 commissioning but not observed.

Ref #	Object	RA/DEC	Calib star RA/DEC	V mag	Obs time	Obs bands
4	IRC+10420	19 26 48.03 +11 21 16.7	19 30 10.83 +11 31 41.3 19 29 04.97 +10 37 50.8	8.5	02:30 ...	J & H broad
6	MWC 297	18 27 39.60 -03 49 52	18 22 24.92 -05 04 48.8 18 26 22.06 -02 53 44.6	9.5	02:30 ...	J, H & K broad
8	U HER	16 25 47.47 +18 53 32.87	16 24 09.08 +16 09 21.9 16 18 49.00 +17 23 58.1	8.5	23:30 ...	J & H broad
10	BD+404124	20 20 28.5 +41 21 51.6	20 22 43.11 +41 59 32.8 20 23 21.47 +40 44 09.1	10.7	03:00 ...	J, K broad & Pa-beta
11	RX BOO	14 24 11.63 +25 42 13.40	14 23 52.27 +27 16 14.7 14 21 12.50 +24 07 59.7	7.97	22:30 ...	J & H broad
12	RS VIR	14 27 16.39 +04 40 41.1	14 28 31.14 +02 47 19.8 14 18 29.82 +03 40 28.3	7	23:00 ...	J & H broad
13	CQ Dra	12 30 06.6576 +69 12 04.045	12 34 44.01 +70 01 18.4 12 47 34.34 +66 47 25.1	4.95	00:00 ...	J & H broad
14	BF Cyg	19 23 53.5 +29 40 29.22	19 22 50.75 +29 39 44.9 19 24 16.35 +27 44 15.2	10.3	02:00 ...	J & H broad

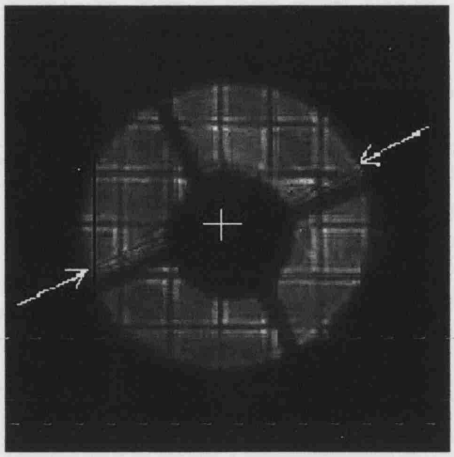

Figure 4.23: Pupil image taken with the 1.6 arcsec mask at the focus. This is the occulting mask that was calculated to be outside the angular alignment tolerances and as is indicated in the image there is a notable shift in the pupil and the telescope vanes are no longer masked.

- Object 1 - Wolf-Rayet star with surrounding nebula structures.
- Object 2 - Luminous Blue Variable (circumstellar dust shell).
- Object 3 - Post AGB star (yellow hypergiant with compact nebula).
- Object 5 - Star with brown dwarf companion.
- Object 7 - Young star with dust-disk.
- Object 9 - Dusty red star.
- Object 15 - Close A2/G0 binary with *speculated* brown (T) dwarf companion at 6".

The object number in the above list is the same as is used in Table 4.5.

On reducing the data, object 3 – HD168625 – has revealed definite structure in the image (Figure 4.24) which compares well with those in the literature, as shown in Figure 4.26. The object was observed using J (300s), H (195s), Fe-II (300s) and Pa-b (120s) filters with the integration times noted, however the nebula structure was only detected in H-band. This structure occurs at about 7 arcsecond from the central star and so would likely have been visible with AO alone.

The reduced data for object 15 – HD150451AB – revealed a faint detection of the T-dwarf companion. Since no field rotations were performed to confirm this was not an AO artefact and the signal to noise for the dwarf was extremely low, no specific conclusions

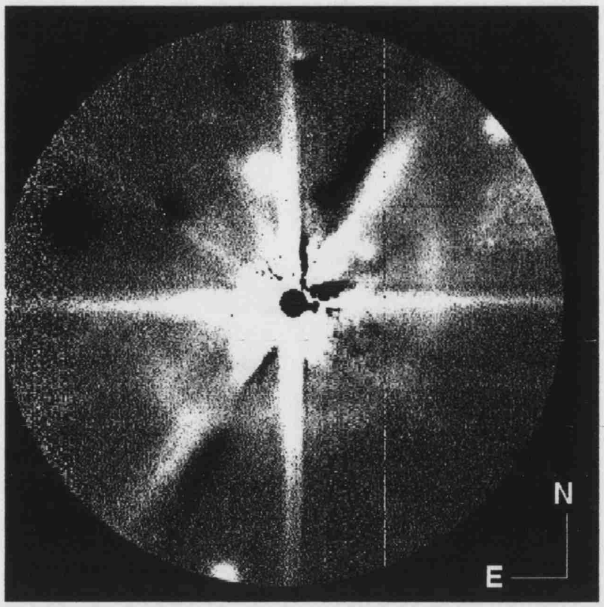

Figure 4.24: H-band image of HD168625 taken using NAOMI+OSCA+INGRID at the WHT. Extra field stars were present in the calibration star

could be made as of May 2002. As of January 2004 there has been a pre-print of data collected over 2 years on this object (Carson *et al.* 2003). This new data in comparison with the measurements made using NAOMI and OSCA on the WHT during the May 2002 commissioning makes it highly likely that the detection made was of the methane brown dwarf HD150451C. The measurements made with OSCA go towards a confirmation of the measurements made by Carson *et al.* (2003). HD150451C is only the fourth known methane dwarf companion to a normal star. For the OSCA data, object 15 was observed in Ksh (260s) and J (200s) filters. Object 15 was the only object that night for which we took a set of sky frames so a good background subtraction was possible. The calibration star used was also good, in that there was not a detectable presence of other field stars (as for object 3 and others) and it subtracted surprisingly well. This is perhaps due to the largest occulting mask (2 arcsec) being used and the stars being well centred behind it. Figure 4.27 shows a final processed image of the T-dwarf. This is a composite of both the J and K data and a matched calibration star PSF has been subtracted. The detection was only just discernable above the noise in J and was slightly more defined in K. Taking the pixel scale in INGRID to be 0.038 arcsec/pixel, the separation of HD150451C from the central star in the image is 6.65 arcsec $\pm$ 0.1 arcsec. The position angle as measured from the central star's north-south axis, in an anti-clockwise direction, is 73 degrees. Estimates

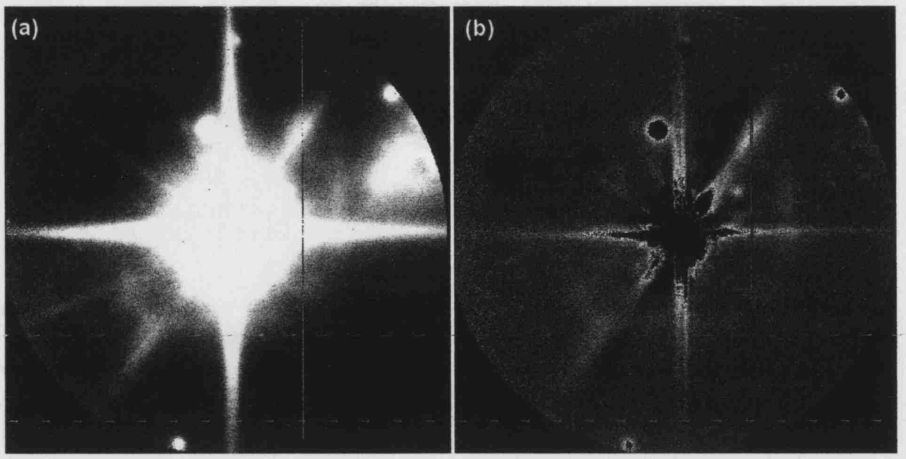

Figure 4.25: Images of HD168625 at different stages of reduction. (a) the image has not has a PSF star subtracted, (b) a calibration star has been subtracted and the values thresholded to remove bright sources.

of the J-K colour magnitude have been made although these are quite rough due to the very low signal. Counts were summed for a 30 × 30 pixel region centred on the detected star in J and K, and the same summation applied for regions the same distance from the central star to get an estimate of noise and errors. This gave an estimate of the J–K magnitude to be 0.65 ± 0.25. Since no infra-red standard stars were observed no estimates of the J or K apparent magnitudes could be made. Carson *et al.* (2003) has measured the K-magnitude to be 16.0 ± 0.2 with J–K as 0.55 ± 0.10, the separation as 6.750 ± 0.015 arcsec and the position angle as 71.0 ± 0.1 degrees. The OSCA measurements, although less accurate, are in close agreement with these values.

No other detections were made of nebulae, dust disks or companions about the other objects described.

In retrospect, too many objects were attempted for observation in the short time available on-sky. The fact that this was our first time in taking astronomical data through a coronagraph, getting used to the requirements of a new instrument in terms of calibration data etc. were other reasons for the lack of astronomical results, although the poor seeing was the overriding factor. However, much experience was gained with the system so that a clear observing guide could be written for future users. The following brief points summarise the requirements for completing an OSCA observation and are taken from the OSCA Observers Guide (Doel *et al.* 2002) produced for the ING as part of the OSCA

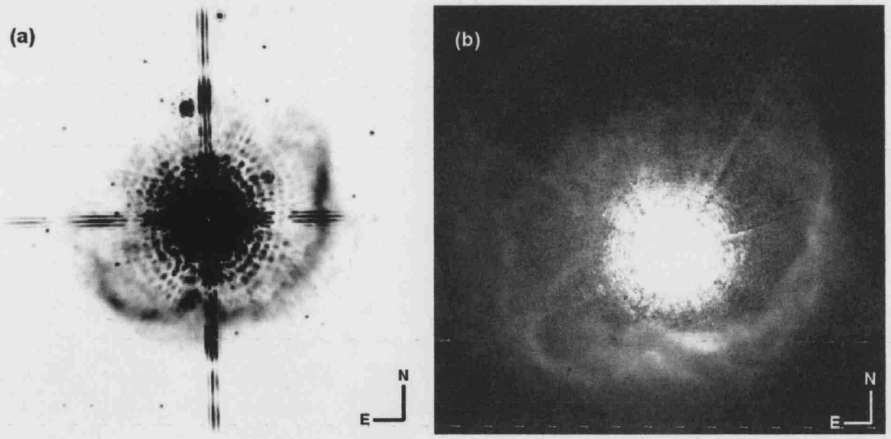

Figure 4.26: Recent images of HD168625. (a) K-band (2.2μm) image taken with the NICMOS coronagraph on Hubble (O'Hara *et al.* 2003). (b) H-alpha image taken with WFPC2 (Pasquali *et al.* 2002)

instrument manuals:

1. Select INGRID filter and cold stop using INGRID GUI.
2. Select the OSCA control panel in the NAOMI GUI.
3. Select the OSCA deployment in button.
4. Focal plane mask sizes can be selected by clicking on the appropriate mask button.
5. Move the telescope until images of the guide star to be used are seen on the NAOMI wavefront sensor.
6. Close loop on NAOMI and move guide star position to centre of focal plane mask using the NAOMI guide probe offsets.
7. Take exposure of object.
8. Open loop on NAOMI and move telescope off object to clear patch of sky.
9. Take a sky exposure (ideally with integration time same as object).
10. Move telescope to calibration star and switch on NAOMI (without any change of settings or movement of pick-off position).
11. Take exposure of subtraction/calibration star.

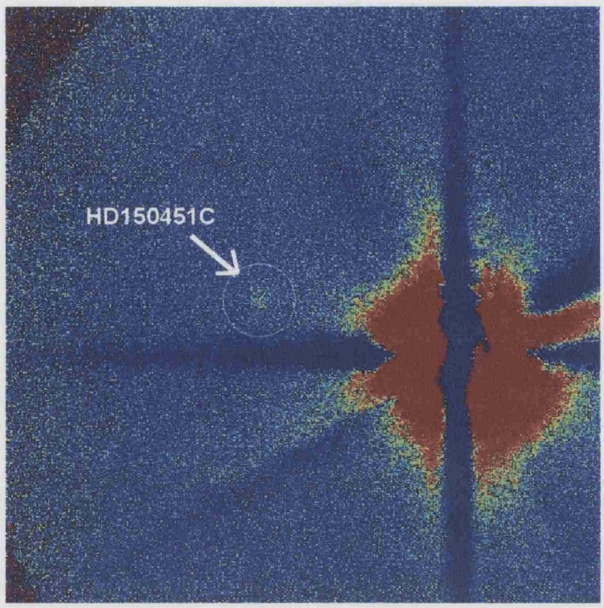

Figure 4.27: Detection of HD150451C (T-dwarf) using NAOMI+OSCA. The central star has had a scaled calibration star PSF subtracted from it to increase contrast. North is up and east is to the left.

12. Open loop on NAOMI and move telescope off calibration star to clear patch of sky.
13. Take sky exposure (ideally with integration time same as calibration star).
14. Open loop on NAOMI and move telescope back to object.
15. Rotate field of view by 60 degrees.
16. Close loop on NAOMI and recentre object on mask, if necessary, using pick-off.
17. Repeat procedures 6-12.

Flatfields for every combination of OSCA and INGRID filters used and INGRID darks must also be taken. Time can be saved if sky images of the same time integration as the object images are taken since these will contain the correct dark current (dark levels do not scale linearly with exposure time) and so the dark current and the sky background can be subtracted in a single step. On the subject of flatfields, a lesson was learned for any future observers of the system: during the observations, many combinations of filters in INGRID were tried with many of the different occulting masks in OSCA and the result was the need to take 16 different flat-fields at the end of the night which is very time

consuming. The advice is to keep combinations of INGRID/OSCA filters to a minimum and to keep using the same occulting mask that best fits your science requirements and the local seeing conditions at the time of observing.

The seeing degraded as the night went on so that accurate centring on the masks again became a problem and performance was adversely affected. It is also noted in many of the images there is a very bright diagonal artefact (bottom left to top right) which caused undesirable subtraction effects. There was only one object in which this was not a problem and the seeing was noted to be very good at this time. It is hence speculated that this artefact is due to the AO having problems with the bad seeing. A number of in-focus ghosts were observed in the images which appear to be present even when OSCA is out of the system (see Figure 4.28). Although not ruling out that some ghosts may be from OSCA it seems likely that most of them are from elsewhere in the system, e.g. the INGRID fore-optics.

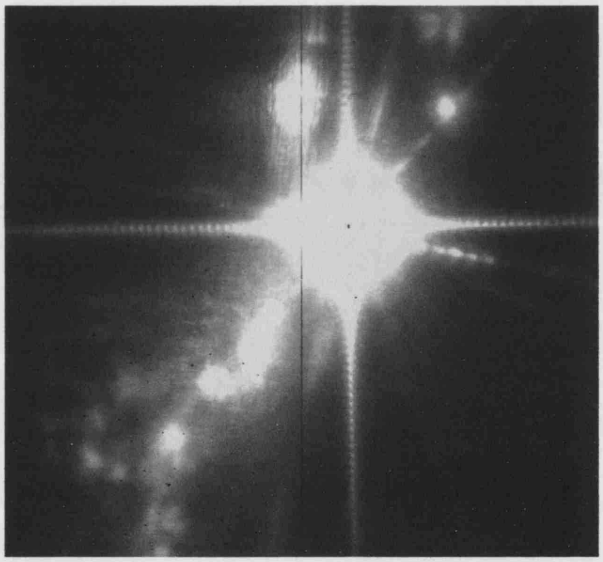

Figure 4.28: H-band, standard star image without OSCA. This is a single star, there is a line of in-focus ghosts from bottom left to top right and there are also some additional out of focus artefacts. The banded diffraction spikes due to the NAOMI square segments are clearly visible.

4.5 Further Commissioning 28 August 2002

Due to the poor weather conditions on the first commissioning run, an additional commissioning run was assigned for two nights over 28-29 August 2002. This was carried out by the staff at the telescope. Prior to the run a list of objectives had been drawn up which included further tests that needed to be done and some suitable science objects to look at if time and conditions allowed. Unfortunately this commissioning run experienced difficulties due to technical problems with the adaptive optics system and so only one night of the on-sky time could be used.

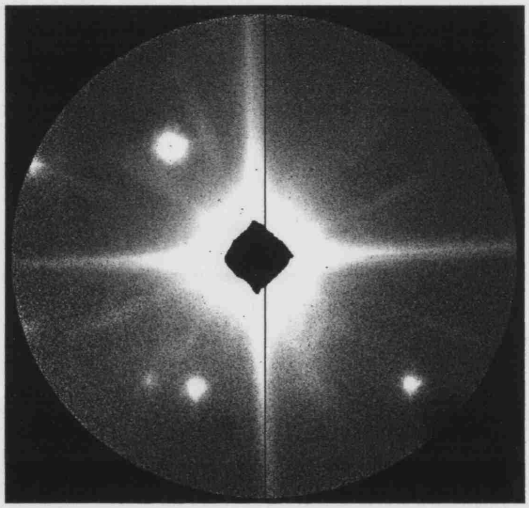

Figure 4.29: H-band image of BD+404124. Counts above 1000 have been thresholded to zero to enhance the faint stars in the open-cluster. Four stars can be seen in the field and a fifth just at the edge (top left).

The list of objects not observed during the May commissioning (Table 4.6) had been left with the ING staff as possible further commissioning/test targets. Of particular interest were MWC297 and BD+404124 which are both Herbig Be stars that are very bright in IR and are part of a cluster (a closed cluster for MWC297, open cluster for BD+404124). The other cluster members are much fainter and so a coronagraphic observation would help detect any faint cluster members close to the bright stars. Due to the short time on sky only one object was looked at, and unfortunately not in the filter bands of interest that were specified. BD+404124 was observed in broadband H filter for a total integration time of 1200 seconds (see Figure 4.29). The system performance during this run is not good -

the AO was having technical problems and it can be deduced from other data that OSCA has not been aligned correctly by the ING (the instruments had been moved off and back to GHRIL between the May commissioning and the August run). The calibration star did not subtract well, so Figure 4.29 is just the summed average of all the object frames, sky subtracted and flatfielded. It has been thresholded so that counts greater than 1000 are set to zero which improves the contrast for the fainter stars that are part of the open cluster.

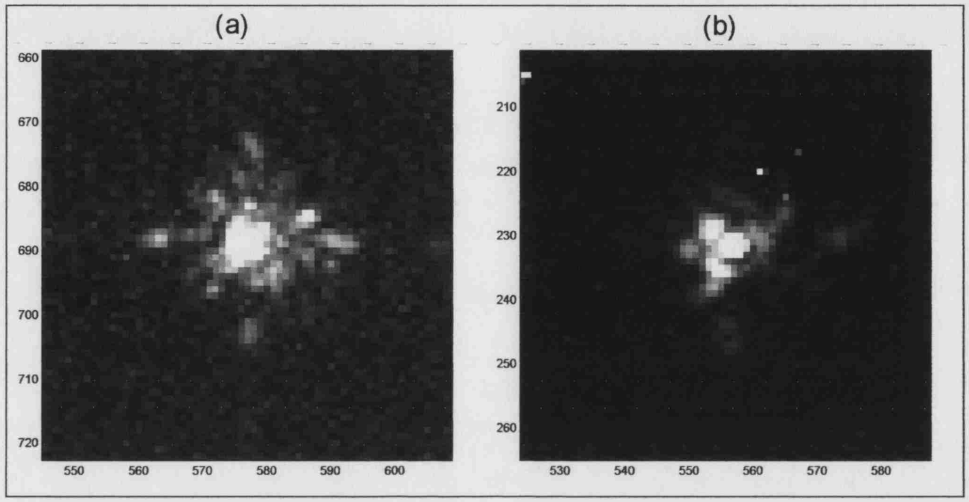

Figure 4.30: H-band images of the pinhole, (a) without OSCA in the beam and (b) with OSCA deployed in the beam. In this case OSCA has been misaligned.

Since OSCA had not been correctly aligned, the data obtained was poor, particularly the shape of the PSF through OSCA. Figure 4.30 shows images of the focus taken on the pinhole just before observing on-sky. The left-hand image is the NAOMI focus without OSCA in the beam, which appears reasonable. The image on the right is the focus after OSCA has been deployed into the beam. Taking into account the loss of throughput due to the Lyot stop, both images have been scaled to the same level (background noise is due to the image falling on different INGRID quadrants). The OSCA focus is badly degraded and as can be seen from the pixel numbers labelling the axis there has been a considerable shift ($\sim$ 460 pixels) in the focus as well. When OSCA is aligned correctly in the beam, the focus should hardly shift on the detector at all (certainly less than 50 pixels). The change in focus quality should be hardly discernable by eye (slightly broader PSF due to the reduced aperture of the Lyot stop) which is not the case in Figure 4.30. The misalignment

of OSCA was confirmed in April 2003 when we returned to fit some upgrades for OSCA. Although the system had been moved to a different bench the internal alignment should not have been changed and it was noticed almost immediately that the top plate of OSCA had been lowered at one end (OAP1 end) so that all the opto-mechanics were sitting on a tilted plate.

4.6 Moving OSCA to GRACE

During 2002 the ING built a new instrument room for the other Nasmyth platform on the WHT. This was installed in December 2002, and all the instruments that were installed at GHRIL were scheduled to be moved over to the new platform, GRACE, during the first half of 2003. GRACE is intended to be a reasonably dust-free environment, especially compared to GHRIL, and more thermally stable to provide the much needed environmental stability in which to operate an adaptive-optics system.

From 8-17 April 2003 we commissioned three new focal plane masks for OSCA, fitted a modified detent arm (to accommodate the thicker substrates of the gaussian masks), upgraded the Lyot stop holder, fitted an anti-scatter mask for the focal plane substrates and completed a full alignment check and installation of OSCA at GRACE. An hour of on-sky measurements were also taken, although the weather conditions were poor and the AO performed sub-optimally.

4.6.1 Commissioning Gaussian focal-plane masks

OSCA was removed from GRACE and taken down to the test-bench in the aluminising area of the WHT. This was not an ideal area for installing the new masks (1 arcsec sharp edged mask and the 2 gaussian masks) due to the dusty environment.

The three new masks were installed successfully using the procedure described in §4.1: a photograph of the apparatus used for this at the telescope is shown in Figure 4.32. The remaining substrates in the wheel were cleaned, as they had become very dusty in GHRIL. All ten positions in the OSCA focal plane wheel were now filled, as detailed in Table 4.7.

Before removing OSCA from GRACE it had been brought to our attention that one or two of the masks had been noted as being smeary on recent images taken with INGRID.

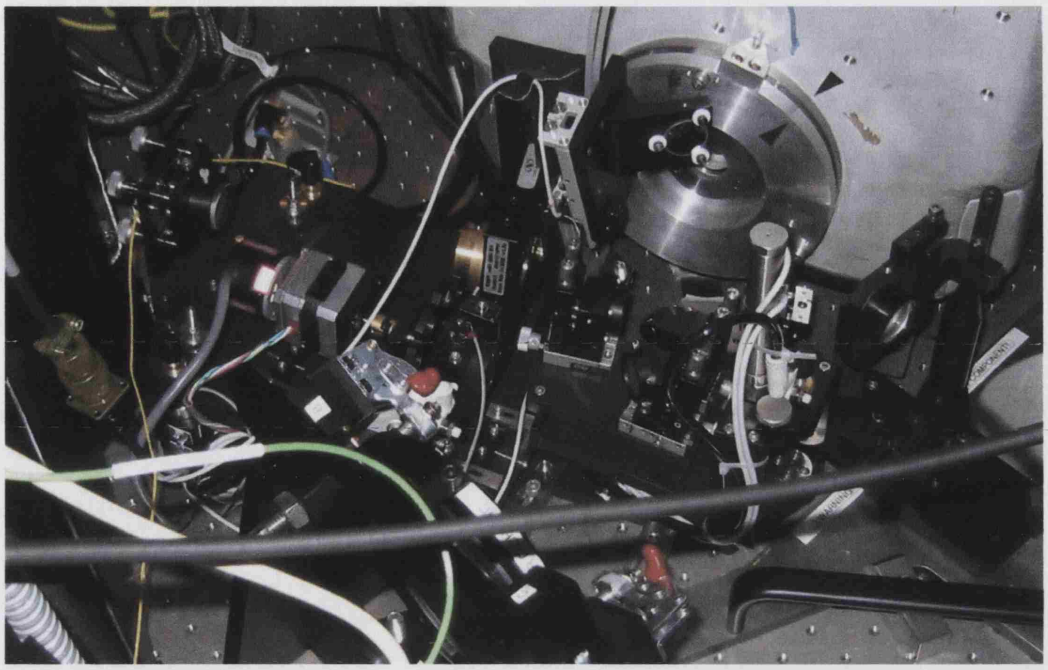

Figure 4.31: OSCA aligned and installed at GRACE. Note the extra optics installed (compared to at GHRIL, Figure 4.18) to accommodate the new integral field unit, OASIS (Figure A.4).

Closer inspection revealed a large and very greasy (perhaps mechanical grease) fingerprint smeared across the 0.4 arcsec mask. A less greasy fingerprint was also present on the 1.6 arcsec mask. Corrosion due to this grease resulted in the 0.4 arcsec mask being washed off completely on cleaning and the 1.6 arcsec mask now has slightly ragged edges. This highlights how important it is that when handling OSCA or working close to the focal plane wheel *the substrate masks must not be touched.*

The accidental creation of a clear position is probably worth keeping as it allows off-mask photometry of a star by simply changing the mask wheel position. This is less time consuming than shifting the star away from the mask using the NAOMI pick-off probe and avoids having to re-centre the star behind the occulting mask again. It also provides a clear position through which the NAOMI simplex procedure can be performed, rather than using the target mask, at the start of the night. This ensures the AO calibration takes into account any static aberrations within the OSCA system to obtain the best PSF on-sky.

To accommodate the new gaussian masks which are thicker than the standard masks, a

Figure 4.32: Optical bench in the Aluminising area of the WHT set-up for the focal mask alignment and gluing procedure. The up-ended table clamped to the bench was an improvisation for a projection screen as the ceiling is extremely high in this area. A picture of the complete set-up is shown in Figure A.2.

modified detent arm was installed. This is based on the same design as the original detent arm except that 2mm has been cut away from the inside edge of the arm that passes in front of the substrates and the screw mounting for the bearing has been swapped to the opposite arm. Note that the main difference between Gaussian A and B is not the FWHM but the profile of the masks with respect to the greyscale levels. Gaussian B has a more extended profile and slightly different grey-level scalings than Gaussian A – on sky testing with OASIS should conclusively reveal which works best. Initial lab tests suggest that Gaussian A gives the better suppression.

A razor-edged anti-scatter mask was also fitted just in front of the focal plane wheel. This successfully masks the edges of the substrates. For a high contrast imaging system it is important to keep scattering to a minimum. Light scattered by the inside edge of the mask substrates, especially for the thicker gaussian substrates, can be reduced by masking the edges with a razor-edged mask which preferentially scatters light backwards and out of the beam. Any chips or damage at the edges of the masks as can be seen in Figure 4.34(a) will also contribute to the scattering and the anti-scatter mask will help reduce the scattering from these too. However the mask does reduce the field of view by about 3

Table 4.7: Final list of masks in focal plane wheel (in an anti-clockwise direction).

Mask name	Additional notes
Gaussian A	FWHM = 0.5"
Gaussian B	FWHM = 0.6"
1.0"	Highest optical density
2.0"	
1.6"	Some damage to edges, alignment not within tolerances
0.8"	
0.65"	
Clear	Previously 0.4" mask
Target	Outer diameter = 1.0"
0.25"	

arcsecs. If this is ever a problem the mask can easily be removed.

After completing the instrument upgrades the internal opto-mechanical components were again aligned using a f/16 laser beam at a height of 15cm. The internal alignment of the OSCA components was checked and adjusted where necessary so that the focus was maintained with OSCA in the beam and with minimum lateral displacement.

4.6.2 Installing improved Lyot stop holder

During the initial commissioning of OSCA, we realised that the design of the Lyot stop holder was not the most practical solution for the job intended. The design of the Lyot stop itself has a 2-way symmetry, and when installing needs careful alignment to ensure it is correctly aligned with the pupil plane of the telescope. The original screw-on and locking ring design for the holder no longer seemed sensible as changing the Lyot stop altered the offset angle and so a re-alignment would be necessary.

A new Lyot stop holder was designed so that mounting the stop on to the SR50 was simply a case of locating a metal post on the mask mounting into a close-fitting hole in the SR50 mounting plate and then locking in place with a screw. This minimises or eliminates the need for any re-alignment of the Lyot stop between stop changes etc.

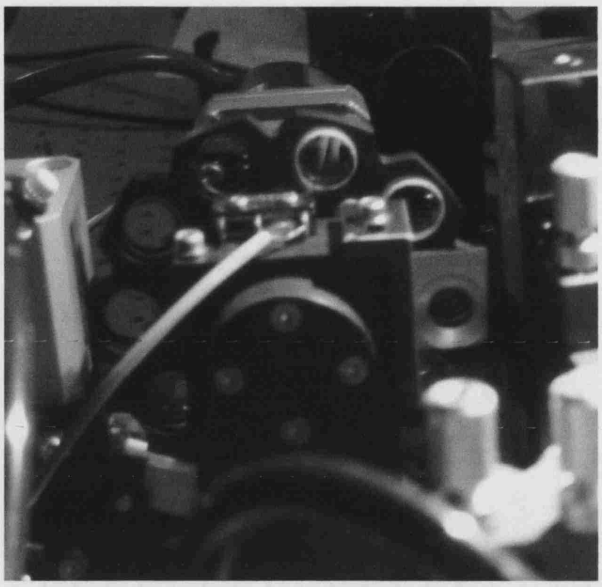

Figure 4.33: Anti-scatter mask installed in front of the focal plane substrate. The two substrates protruding from the wheel are the newly installed gaussian masks.

OSCA was taken back up to GRACE and re-aligned on the bench using the procedures previously described. In an attempt to move the focal plane masks away from the INGRID quadrant boundaries it was noticed that the field had a gradient across it due to vignetting. OSCA was designed for the focus to be on axis and shifting the focus in this way will vignette the beam and cause the pupil to shift by an amount dependent on the amount the focus is shifted. The pupil may then need to be re-aligned and other undesirable effects might be created that have not been anticipated. It is recommended that the OSCA masks are kept as close to on-axis as possible. Problems with objects of interest falling into the detector quadrant boundaries can be avoided by using field rotations between observations. This technique is also necessary to confirm detections using an AO coronagraph.

The co-location of the Lyot mask with it rotation axis was checked and found to need adjustment. This was carefully rectified by measuring the centre of rotation of the Lyot mask holder on the SR50 and positioning the centre of the Lyot mask over this using a micrometer xy-stage and a magnified video camera + TV. This took longer than expected since the new Lyot holder does not have much leeway for positioning the mask on the holder, so there was no remaining time to align the Lyot stop to the telescope vanes (i.e. to find the offset angle that needs to be entered into the control software).

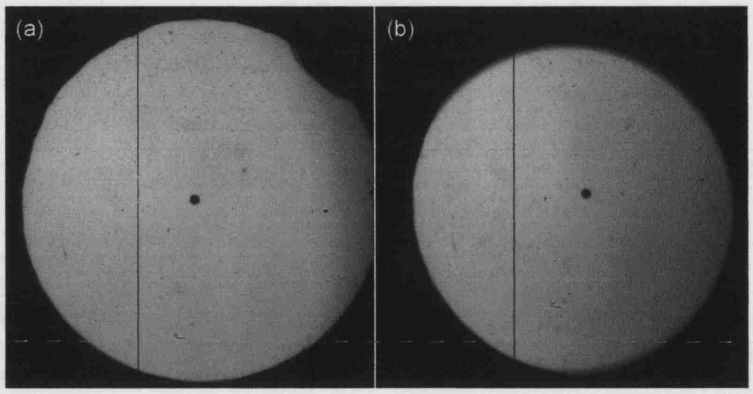

Figure 4.34: Flatfields taken before and after installing the anti-scatter mask

The Lyot stop alignment with the telescope vanes was completed in July 2003 prior to a scheduled observing run with OSCA. Testing revealed that the formula used to control the Lyot rotation at GHRIL (Equation 4.3) does not apply at GRACE. The problem occurred due to the query over the sign of the derotator angle. The information had been that the derotator always gave a positive angle regardless of the direction of rotation, however the derotator angle is in fact positive in a clockwise direction and negative in the anti-clockwise direction.

$$\text{Lyot angle} = 360 - (-2 \times (\text{Derotator angle}) - \text{Zenith angle} - \text{offset}) \tag{4.4}$$

A few combinations of formulae were tested before Equation 4.4 was found to be correct.

4.6.3 On-sky testing

The night of the 15 April was subjected to bad seeing conditions. Although the OSCA masks were well centred over the star, no significant amounts of suppression were recorded. Calculations with a simulated NAOMI PSF give a Strehl estimate of approximately 2% and Figure 4.36(a) indicates that the mirror segments were probably not co-phased. OSCA will perform extremely poorly in such a situation.

The gaussian masks were tested to check their usability in the NIR (H-Z bands) but were found to be unsuitable (the ND levels were manufactured to function correctly only for wavelengths between 0.4 - 0.8μm).

The night of the 16 April had improved seeing conditions, but there was 100% cloud cover

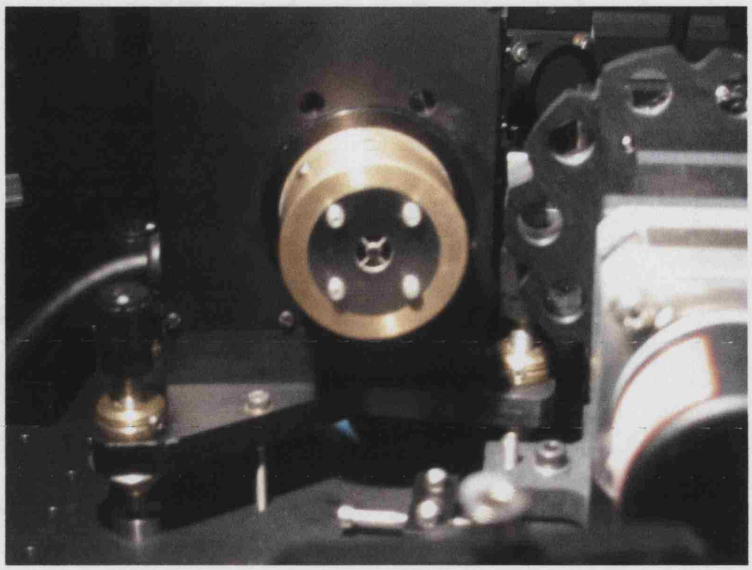

Figure 4.35: Photograph showing the improved Lyot stop holder attached to the SR50 rotation stage.

for a significant portion of the night with only a small clearing appearing near dawn during which the OSCA measurements were taken. Unfortunately the current best method of centring the OSCA masks over the star is by using a special zoomed in, real-time INGRID window from the NAOMI GUI. The NAOMI operators on this night were unfamiliar with this function and consequently the masks were poorly centred over the star and in some cases the centre of the star missed the masks completely. Ideally an automatic routine would be written to centre the masks over the star to avoid similar errors in the future. The Strehl ratio achieved on this night was approximately 4% and the mirror appears to be co-phasing, although Figure 4.36(b) shows that two faulty segments are causing problems. These were identified as segments which have the strain-gauge feedback switched off.

4.6.4 Suppression tests using the artificial star

A pre-NAOMI filter mechanism (Figure 4.37) was installed at GRACE during May 2003. This provides a choice of different sized pinholes (with a new, small 2μm diameter pinhole) to create the artificial star which a number of system alignments and calibrations require. Unlike the old pinholes these are mounted on a motor driven linear stage in x (across the beam) with a manual micrometer adjustment in y. If the pre-AO occulting mask is

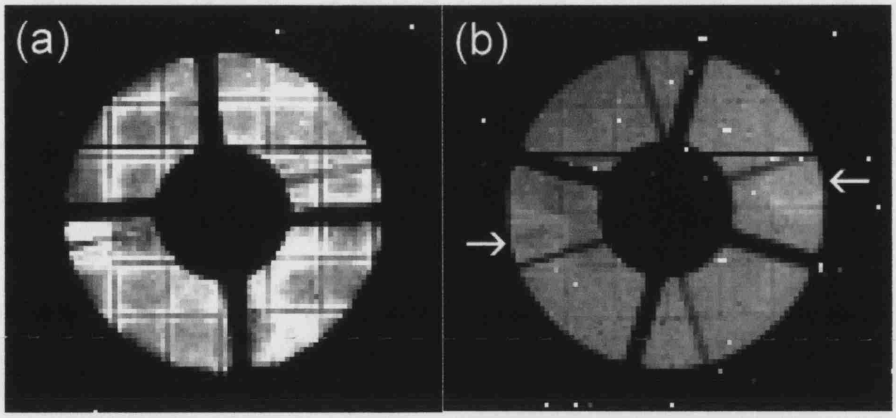

Figure 4.36: Images of the pupil plane taken during the April run in different seeing conditions. (a) 15 April, very bad seeing - the NAOMI pupil appears to be out of phase; (b) 16 April, better seeing conditions - the NAOMI pupil appears to be co-phasing although two of the faulty segments indicated can be seen to be out of phase.

manufactured it will be installed in an empty slot in this device.

The provision of a fully mobile artificial star finally made it possible for suppression testing to be done with OSCA off-sky. During the afternoon set up before an observing run in July 2003 a set of data was taken using all of the OSCA occulting masks except the gaussians in H-band using the 2μm pinhole artificial star. The OSCA Lyot stop was not used for this, the OSCA_OBS pupil stop in INGRID was used instead. This was because the OSCA system had been discovered to be misaligned once again (pupils not aligned) and there also appeared to be a slight misalignment with the telescope/NAOMI pupil and INGRID. This had occurred when all the instruments were moved off the bench between April and July to allow the installation of OASIS in GRACE. There was no time to do a system alignment, so to minimise problems the OSCA Lyot stop had been removed.

The results from the artificial star are plotted in Figure 4.38. All frames were dark subtracted and calibrated to the same integration time. The image taken with no mask in place was very short to avoid saturation of the peak and so becomes limited by read-out noise at low counts which occurs in the wings of the PSF. The artificial star was not moved between occulting mask changes. The FWHM of the artificial star is 2.5 pixels ($\sim$0.1 arcsec) and the Strehl ratio is approximately 55% - based on a comparison with a

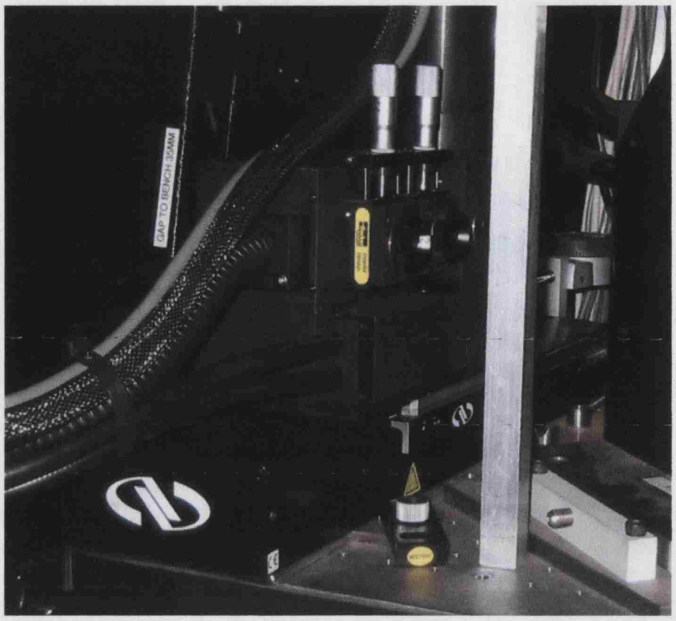

Figure 4.37: The new pre-NAOMI filter mechanism. The different filters are placed on a sliding mechanism for automated change and there is also a y micrometer manual adjustment. This provides a selection of various size pinholes to create an artificial star, needed for calibration and testing. The OSCA pre-AO dichroic focal mask will also be installed here when completed.

simulated PSF. An image of the PSF taken through the OSCA clear position is shown in Figure 4.39. The suppression is poor due to the misalignment of the pupils which reduces the effect of the pupil stop in INGRID.

Some additional data was taken with the gaussian occulting masks in Z band, (the shortest waveband possible on INGRID) as shown in Figure 4.40. Although there is some suppression of the peak of the star there is no subsequent suppression in the wings. This is the best that can be expected at Z-band ($\lambda = 1\mu$m) given the change in ND-levels with wavelength (Figure 2.13). The maximum wavelength at which these gaussian masks should be used is 800nm, for which the peak ND value is 2.5. At 1000nm, this peak ND level has dropped to 1.2, although the graph in Figure 4.40 indicates an ND value of 0.9 at the centre, which suggests a slight mis-centring of the masks over the pinhole. No suppression can be expected in such a situation. Proper testing of the gaussian masks must await the availability of a visible light camera.

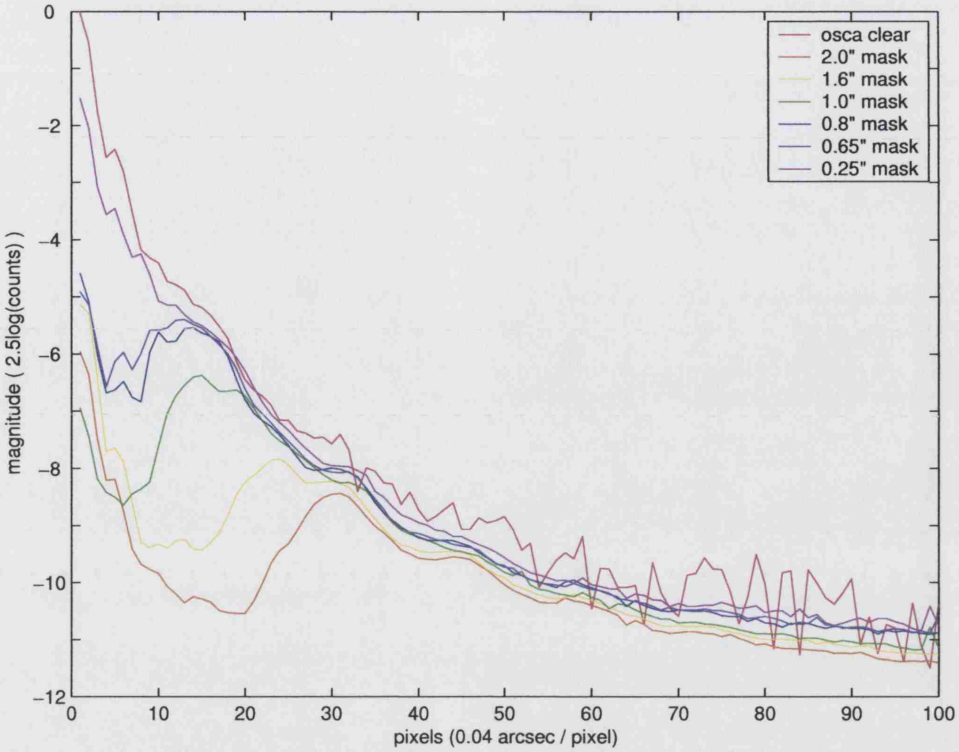

Figure 4.38: Suppression curves taken using the artificial star at the WHT and the various occulting masks on OSCA.

4.7 Observing time with OSCA

Numerous science studies can be developed with the use of a coronagraph and an adaptive optics system, as mentioned in §1.4. An interesting science case was proposed by W. F. Thi, a visiting astronomer at UCL, for using OSCA with NAOMI in the search for large planet-type objects in protoplanetary disks. The observation plan was to determine for the first time the location of the dust *and* gas in the disk around the Herbig-AeBe star HD141569. HD141569 is already known to possess a disk with an inner dust gap. This can be attributed to signs of planetary formation such as a giant planet in orbit which has swept up the disk material, however other theories exist describing gravitational interactions between the other stars in the system to account for the banded disk structure. A test for the planet model is the presence or absence of a gap in the gas. A gap for both dust and gas would strongly support the theory that a proto-planet is in orbit about HD141569. For these observations the disk around HD141569 needed to be observed in both broad (J and H) band and narrow (H_2 emission lines) band filters. Since the detection of a gas-gap was

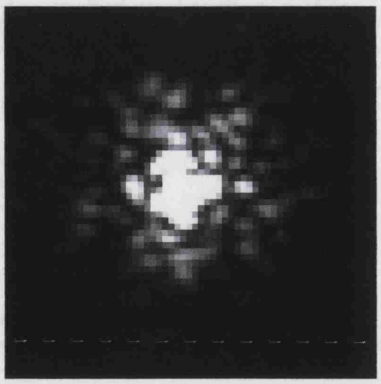

Figure 4.39: Image of the artificial star taken through the OSCA clear position.

of particular importance a deep integration was indicated for the narrow bands centred around the molecular hydrogen ($v = 1 - 0$ and $v = 2 - 1$) emission lines. The author (SJT) was named as a collaborator on this project and was the only observer.

A few months before the observing run, a paper was published with the first ground based, near-IR observations of HD141569 with an AO system and coronagraph (Boccaletti *et al.* 2003). Previous images of HD141569 had only been taken with the HST (Weinberger *et al.* 1999). The results of Boccaletti *et al.* (2003) (see Figure 4.41) did not give confidence that a detection would be made with the system at the WHT. As can be seen from the

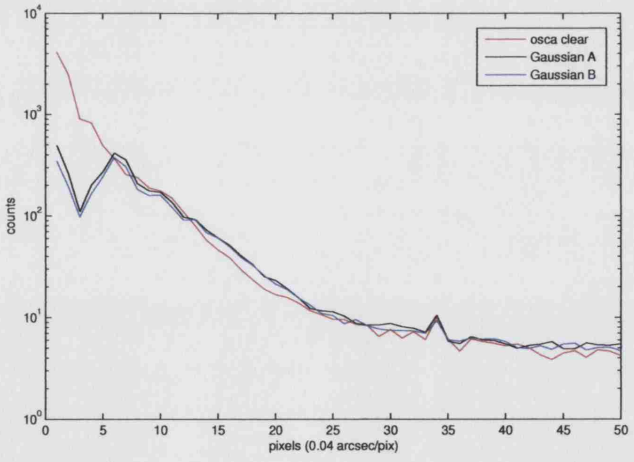

Figure 4.40: Radial averages of the artificial star behind the gaussian masks in Z-band.

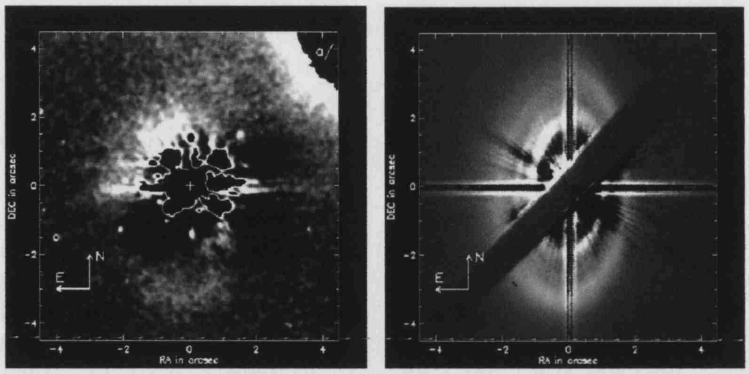

Figure 4.41: Ground-based, AO image (left) (Boccaletti *et al.* 2003) and Hubble/NICMOS image (right) of HD141569 (Mouillet *et al.* 2001).

figure, the image is rather poor although it is taken with the 5m Palomar telescope using PALAO - a 241 actuator adaptive optics systems at Cassegrain. This has achieved excellent results in the past but due to poor conditions during these observations only obtained a Strehl ratio of 16%. The WHT is a 4.2m telescope and NAOMI is a lower order AO system, that usually obtains a 10% Strehl ratio at best. However since NAOMI is a segmented mirror and the Palomar is a continuous face-sheet mirror the two Strehl ratios cannot be compared directly. If the segmented mirror grid is taken out of the simulated PSF and this used to calculate the NAOMI on-sky PSF, the Strehl ratio is 3% lower (§3), showing that a segmented AO system is lower contrast by design. Given this information the likelihood of obtaining the planned detections was low.

The observing nights were scheduled much later in the year than the ideal dates requested in the proposal. The preferred dates were mid-May to mid-June when HD141569 is up for most of the night and reaches its maximum altitude attainable from La Palma. The two observation nights were given in late July: at this time HD141569 was only above the minimum altitude for AO ($\sim 40°$) for half the night. Due to this the observations were taken over 2 half-nights and one full-night. About half a night was lost due to technical problems with the filter change mechanism in the INGRID detector. For the half-night in which HD141569 had set an additional object, HD210277 was observed. An unconfirmed observation of a disk (Trilling *et al.* 2000) had been made around this object in H-band and is shown in Figure 4.42.

Instrument checks before the observing run revealed that the pupils were not aligned and

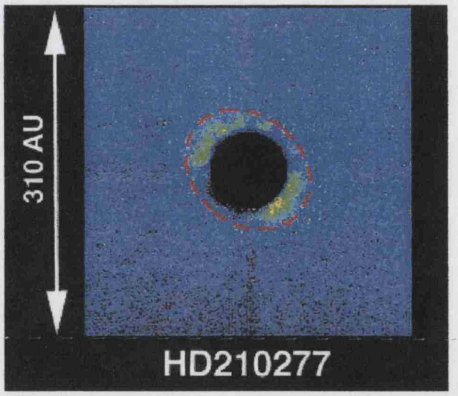

Figure 4.42: H-band image of the circumstellar disk around HD210277 as detected by Trilling *et al.* (2000). This is a 2-minute exposure (selected from a total of 32 minutes) taken with CoCo, the cold coronagraph on IRTF. North is up and east is left.

that the OSCA Lyot stop contributes too much thermal noise in K-band. For these reasons the OSCA Lyot stop was removed and only the cold stop in INGRID used.

The revised observing plan was as follows:

- Full-night - INGRID filter mechanism fault, HD141569 data discarded.
  HD210277: 2×20 minutes (H-barr) at two different sky rotations.
- Half-night 1 - HD141569: 60 minutes (H_2 v=2-1), 30 minutes (J band).
- Half-night 2 - HD141569: 60 minutes (Kcont), 15 minutes (J band).

Sky frames were taken for all data (every 10 minutes in K and every 15 minutes in J) to allow more accurate background subtractions. Two different PSF subtraction stars were observed for each target, although with much longer intervals than is ideal due to time constraints. It became apparent that the PSF stability was poor so that subtractions across timescales greater than 5 minutes were contaminated with many AO residuals. Figure 4.43 shows this PSF change over time, taken from J-band images of the HD141569 data.

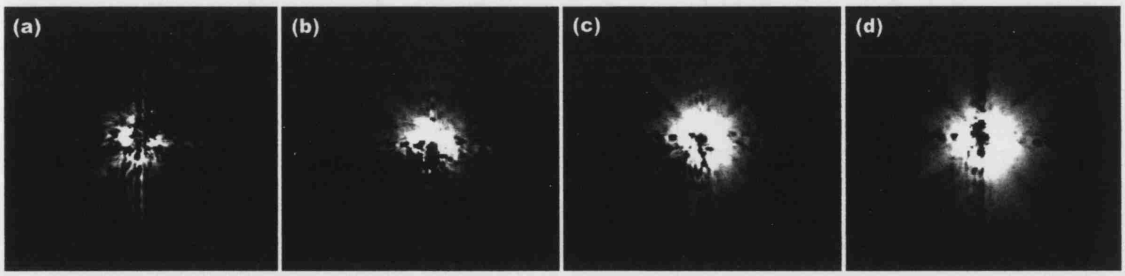

Figure 4.43: A series of same object subtractions over time in J-band to illustrate the changing PSF due to NAOMI. (a) after 1 minute, (b) 5 minutes, (c) 15 minutes and (d) 40 minutes. Atmospheric seeing conditions were average (measured as 0.7 arcsec in V) and each of the images used in the subtraction had an integration time of 30 seconds. All images are scaled to same level.

4.7.1 OSCA data reduction

The data reduction procedure for OSCA is similar to that for most astronomical data with a few additions. All data should be dark subtracted and flatfielded. Background sky levels should also be subtracted, however this cannot be accurately estimated in the small field of view so separate sky frames must be taken. If the exposure of these is chosen to match the target integrations then the dark current and background can be subtracted in one operation. If the centre of the star is visible through the mask or other stars are present in the image these should be used to accurately add the object frames together. Images added together that differ in star-mask centring by just a few pixels will not produce a good PSF subtracted image. If this is required, these images should be left out of the stack. Similarly, a PSF star that is centred behind the mask differently from the target star will also result in a poor subtraction. These are the primary reasons why an automated mask centring routine is required; human judgement in this matter is not accurate and is time consuming.

If the observer had time and the PSF star was observed at regular intervals between the target observations, the data should be reduced in sets to obtain the best PSF subtraction at each stage and then the sets of subtracted images added to complete the reduction. To best achieve a PSF subtraction the centre of the star in each must be measured and a radial average performed about this point. The PSF star must be scaled so that the wings of the star (the profile just outside the occulting mask) closely match the level of the target

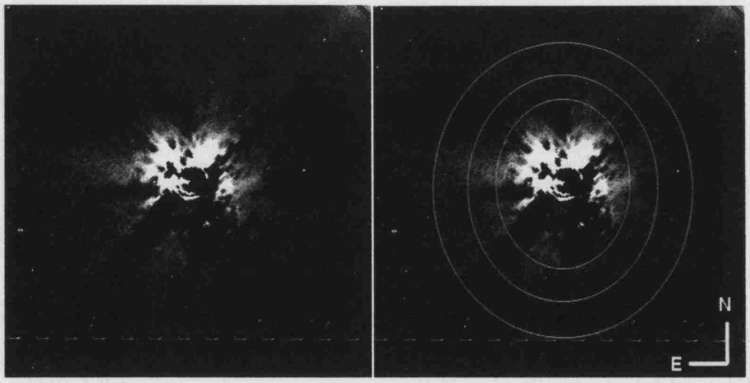

Figure 4.44: J-band image of HD141569 taken at the WHT using NAOMI and OSCA. The image on the right has had the approximate outline of the disk features superimposed upon it for comparison - no parts of the disk are obvious.

star. The two frames can then be subtracted using any offsets determined from the centre measurements. The scaling of the PSF may need to be adjusted by trial and error to get the desired result. After these stages are completed zero-masks can be applied to the images to black out the part of the image under the mask, INGRID quadrant boundaries and the non-OSCA field of view.

4.7.2 Results

Analysis of the HD141569 images shows no evidence of the disk, both in broad band J (see Figure 4.44) and in the hydrogen emission lines. The images were analysed both with and without a PSF subtraction, since PSF subtraction can leave misleading residuals.

The data for HD210277 was reduced and analysed in a similar manner but no evidence for a disk could be detected here either.

Some discretionary time (30 minutes) was awarded for time lost during the April commissioning. The target MWC297 was picked from the May 2002 commissioning list (see Table 4.6) given the interesting science case put forward by J. Drew and M. Barlow and based on the work of Testi *et al.* (1998). MWC297 is a very IR-bright Herbig Be star and the data presented by Testi *et al.* (1998) suggests a compact cluster of stars around MWC297 out to a radius of about 20 arcseconds. The next brightest stars in this cluster are estimated to be approximately 1000 times fainter in K than the central star so high

resolution, coronagraphic imaging of this object could reveal the underlying cluster.

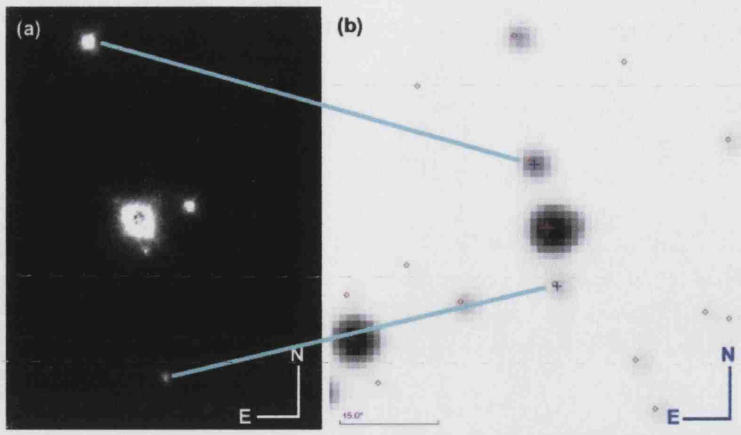

Figure 4.45: K-sh filter band image of BD-04 4476 (SAO 142339). (a) There are 4 stars about the central (occulted) star at distances of 1.59, 2.72, 8.26 and 9.67 arcseconds. The K-magnitude of the closest star is 12.2 ± 0.3. Estimates on the occulted star are difficult due to the mask, but very roughly puts the K-magnitude at around 8.9. (b) 2MASS image of same region. The two stars marked with blue crosses are the outermost stars in the OSCA image.

Due to the limited time, nothing of interest was uncovered in the MWC297 data. However the PSF star chosen for MWC297 revealed a group of 4 stars in close proximity (see Figure 4.45(a)), with the closest being only 1.6 arcseconds from the central PSF star. No off-mask images were taken for this object since it was originally only to be used as a PSF subtraction star, so estimates of the K-magnitude are difficult. The occulted profile was compared to that of an un-masked IR standard star and scaled so that the wings closely matched. This scale factor (corrected for integration time) was then used directly with the IR standard star magnitude and applied to the usual magnitude-flux relation to obtain a value of 8.9. The K-magnitude estimate for the closest star is 12.2 ± 0.3 – a value for the flux contribution from the central star has been subtracted and was estimated by plotting radial averages about the star. The central star and the outermost two stars in this image are listed in the latest 2MASS catalogue of point sources. This lists BD-04 4476 as having a K magnitude of 8.5 and the outer two stars 12.6 and 11.3 respectively. These agree with our measurements when we take into account the estimates involved with the focal stop and the fact that 2MASS cannot resolve the inner two stars, so their flux is added to the

central star. However few studies have been done on this central star, classified as a B3 star but without a luminosity class and efforts to retrieve a spectrum for the star were fruitless. It appears that no spectra (or no good spectra) have been taken for this star and there even seems to be some debate on the V magnitude, with one source quoting 8.8 and another 10.2 (Knude 1992). Assuming the star has been classified correctly, and given our estimates on the K magnitude, 8.8 is the more reasonable of these two estimates. The parallax for the star is given as 10.66 $\pm$ 2.15 mas, putting its distance at 93.8pc. It is possible that the nearby star (1.6" separation, $\sim$150 AU) is a low-mass companion (red-dwarf) to the central B3 star, however more data in different wavebands and evidence of orbital motion would be required to support this. A service proposal has been put in to obtain data in Ksh, J and H on this object. This will help determine any J–H, H–K colour relation which will help constrain the classification of the possible companion, e.g. Wachter *et al.* (2003). Deeper observations will not only improve the S/N ratio but may reveal more stars that would indicate a potential compact cluster.

CHAPTER 5

Discussion

5.1 Review

Overall, this project has succeeded in its aims to produce a high precision stellar coronagraph on a low budget, over a short timescale and meeting all the design specifications. However, due to the way in which commissioning programs are implemented at the WHT, a complete and thorough testing of OSCA has not been possible and as a result there have been no performance tests done with OSCA in optimum seeing conditions *and* optimum alignment (both of OSCA and other instrumentation). A good substitute for this optimum condition set-up would have been to test OSCA on a lab-based AO system with a turbulence generator, however these facilities were not available. It is therefore very difficult to make comparisons between the system that was simulated and the results obtained on the actual telescope. As an example, the magnitude suppression 0.5 arcseconds from the edge of the 1.0 arcsecond mask at a wavelength of 1.6μm in the original simulations and on the artificial star is 1.85 and 0.29 respectively. The measured value is a factor of 6 worse than the prediction. The pupil planes were poorly aligned for this measurement on the artificial star which reduces the Lyot masking and although the artificial star Strehl ratio is more comparable to the simulations than those obtained on sky, any system problems will significantly affect performance. The best on-sky suppression result (taken during the first commissioning run) in H-band using the 2.0 arcsecond mask was a magnitude suppression of 0.75, 0.5 arcsecond from the edge of the mask, compared to $\sim$ 1.9 in simu-

lations. Although the seeing was bad during these observations the system was very well aligned and the AO was performing well given the conditions. The result is a factor of 2.5 drop in system performance compared to simulations. Again this shows how crucial it is to have a well aligned system.

The mechanics and electronics of OSCA have not failed to date and succeed in maintaining the required positioning and alignment, including that for the Lyot stop rotation as the telescope tracks. The deployment mechanism for OSCA has proven to be very successful; it allows OSCA to be raised in to (and out of) the beam very quickly and with consistently accurate positioning. The main problems have been in the alignment of OSCA by the telescope staff – both the internal alignment and the exact positioning of OSCA with respect to NAOMI and INGRID on the optical table. OSCA requires an extremely accurate alignment which was perhaps unfamiliar for those maintaining it and some of the adjustments are awkward and difficult to perform on a bench that is becoming increasingly crowded with instrumentation. The performance of OSCA also strongly depends on system-wide alignments such as all the pupils (WHT, NAOMI, OSCA, INGRID) coinciding to within 5% of each other: this was rarely the case during observation runs. These alignment problems were exacerbated by the continual removal of the different instruments from GHRIL/GRACE, e.g. INGRID is also used as a detector at the Cassegrain focus.

The performance of OSCA also depends on the performance of the adaptive optics system, NAOMI. This was another area in which there seemed to be a number of difficulties and problems. Initially NAOMI was installed at GRHIL which is a non temperature controlled and dusty environment. The temperature variation over the course of a night caused particular problems for the adaptive optics due to the temperature sensitivity of the strain-gauges on the actuators. This would result in poor stability of the system and a worsening of the correction over the course of the night. NAOMI is also known to have a number of faulty segments on the correction mirror where the strain-gauge feedback has been turned off. Two of these segments are within the OSCA pupil. The nature of the fault usually results in these segments not phasing with the rest of the mirror, causing unwanted artefacts in the pupil and reducing the effectiveness of the Lyot masking. There are also problems with the flattening of the adaptive mirror at the start of the night which means the initial quality of the PSF can vary by a significant amount between consecutive set-ups. This problem has been reduced to some extent by the move to the temperature

controlled laboratory at GRACE. The Strehl ratio obtained by NAOMI is worse than expected, $\sim$ 10% at best compared to $\sim$ 40% in the initial simulations at H-band. The performance of OSCA is therefore degraded since the amount of light removed by the focal plane stop is reduced and the amount of light in the AO halo is much larger, decreasing the image contrast. This is particularly noticeable when trying to observe faint, diffuse structures close to stars such as circumstellar dust disks. These problems are gradually being addressed, but over a longer time-scale than was originally anticipated. As more staff become familiar and trained with using AO systems, optimum operation of the system over an observing run is also more likely. However, the faulty segments, large levels of scattered light from these and the gaps/edges of the segments can only be fully resolved by a new adaptive mirror.

An unexpected problem encountered when using NAOMI and OSCA was saturation of the wavefront sensor (WFS). For astronomy observations a maximum guide star magnitude of $\sim$ 14 is given, however it does not state a lower (brightest) limit. This has proven problematic for several OSCA science targets since coronagraphic observations are generally on bright objects, that must also be used for the wavefront sensing. For targets with a magnitude lower than 5 (i.e. brighter), the WFS saturated and subsequent light spillage between the quad cells rendered them indistinct from one another. This has removed the ability to observe interesting stars such as P-Cygni which has a circumstellar dust shell. A neutral density filter of suitable quality should be commissioned in a wheel/slide interchange prior to the WFS to overcome this restriction.

Included in these AO improvements should be the creation of software to automatically centre the corrected star behind the OSCA occulting masks. For future successful use of the coronagraph this should be made a priority. Poor centring of the star, as shown in previous sections, degrades the OSCA performance significantly and makes data reduction and analysis difficult. Even a small difference in centring between objects renders the PSF subtraction star process ineffective due to the asymmetry in the wings of the stars. The centring process is also very time consuming and wastes valuable on-sky time. Human judgement in placing the star PSF is neither consistent nor accurate. Observing with OSCA would become much simpler and more time efficient if this human error was removed.

As mentioned previously, some of the adjustment mechanisms for aligning OSCA are very

awkward to use and make it difficult for many people (especially those with large fingers!) to align it. Perhaps the most precarious of these is the Lyot mask positioning. The rotation stage (to which the Lyot mask is attached) is seated upon a 3-point mounted platform so that the only adjustment possible for correct positioning of the mask is by turning the 3 locking screws, either one at a time to generate a tilt that gives a sideways adjustment or all 3 together for height. During these adjustments the entire Lyot assembly is relatively easy to topple over as the screws are quite stiff and sometimes slip in the mount. One of these screws is uncomfortably close to the focal plane masks so extreme care is required not to knock against them and cause damage to the delicate mask surface. If the opportunity arose to change this unit, it would be best to implement a design in which the SR50 plus Lyot mask were seated on a platform that could be height adjusted from the front using a single, locking thumb screw. If this was also seated on a small slide mounting (similar to those used for the flat mirrors in OSCA), easy adjustments to the x-positioning (across the beam) would be possible using only one screw on the far side of the instrument from the focal masks.

OSCA does not appear to contribute to any ghosting in the system (the ghosting present in images is from the INGRID fore-optics and possibly the derotator optics as well). The use of wedged substrates for mounting the focal plane masks has therefore been successful in avoiding this problem. Other coronagraphic systems (e.g. CIAO) initially had problems of this nature. Also, when properly aligned, OSCA does not noticeably alter the PSF. The only change is that with the Lyot stop in place, the PSF is broadened as expected due to the reduced aperture size. Thermal emissions from OSCA, most notably in K, have not been quantified and neither has the loss in performance due to this. In K-band the masks can be seen to glow in sky images and it is best to remove the OSCA Lyot stop when observing in K to reduce the additional thermal background that this introduces. In hindsight, a change to the optics would have been to use sapphire for the focal mask substrates rather than Spectrosil WF. Although sapphire is slightly more expensive and difficult to obtain with high surface accuracy ($\sim \lambda/10$ instead of $\lambda/20$) due to the hardness, the wedge angle required and the small size of the windows, it is likely that the high durability of sapphire would have provided the better surface in the long term. Considerable damage was done to the Spectrosil windows during the photolithography process (surface scratches, chipping from edges, general surface deterioration from the entire chemical process); sapphire, being

very hard, is much more resistant to scratching and would have maintained a better surface quality for the finished product.

The OSCA focal plane masks have a higher transmission in the IR than expected. However, the 1.0 arcsecond mask which was manufactured after initial commissioning, benefited from data obtained on the original masks' performance and as a result was made with a very thick chromium deposit and accordingly a much lower transmission. This mask is the best one to use with the 2.0 arcsecond mask as second choice to obtain the greatest suppression. It was unfortunate that two of the focal masks were damaged at some point between commissionings. Considering the performance of the AO, the 0.25 arcsecond mask is perhaps too small to provide any useful suppression. The 0.4 arcsecond mask which was destroyed would therefore have been the ideal one to try during optimum observing conditions in order to see closer in to a bright star. It is advisable not to use the 1.6 arcsecond mask as it was partially damaged and so has questionable performance. Since OASIS has not been made available for use with OSCA yet, it has not been possible to use the visible light camera. Due to this, no testing of the gaussian masks at their specified wavelength range has been possible. The laboratory experiments with these masks at 600nm have shown good suppression, especially at close proximity (compared to similar hard-edged masks) to the central source.

As with all precision optics, it is important to keep OSCA very clean to avoid the extra light scattering caused by dust. Although a dust cover was planned for OSCA, possibly with air-flow to over pressurise and prevent contaminants from entering the system, this was never implemented. Some of the very fine particulate dust that was picked up at GHRIL has not been removable, since more thorough cleaning methods cannot be employed on the delicate optical surfaces (over-coated silver and focal masks). GRACE is designed to be a clean environment so a dust cover should no longer be necessary. In addition the GRACE optical bench is so full of instrumentation that a cover for OSCA would probably not fit at present.

A number of interesting astronomy targets have been observed during the OSCA commissioning runs. The most positive results were the very faint detection of the T-dwarf in HD150451 and the discovery of a potential close companion ($\sim$ 150AU) in BD-04 4476. More data is required to draw any firm conclusions on the nature of these objects.

The problems discussed in this section highlight the benefit in having a fully integrated, cooled AO + coronagraph + detector system.

5.2 Future developments at the WHT

There are several plans in the pipeline for upgrades and new instrumentation at the WHT. The plans to upgrade NAOMI to a laser guide-star system have been approved. Tests are planned to trial a continuous face-sheet mirror in place of the segmented mirror. Whether the current adaptive mirror will be permanently replaced by this will depend on the outcome of these tests. Approximately 10% of the NOAMI segments are faulty and cannot be repaired and as discussed in previous sections a segmented mirror design is best avoided when the aim is for high contrast imaging. An upgrade to a continuous face-sheet deformable mirror would be a great improvement, especially for the performance of OSCA.

5.2.1 OASIS

Although OASIS was commissioned at the telescope in April 2003 it cannot be currently used in conjunction with OSCA. The bench at GRACE has become very overcrowded and OAP2 in OSCA impinges upon an OASIS fore-optic when deployed. Due to a change in the design after OSCA was commissioned the exit beam from OSCA is not in the pick-up range of the OASIS optics and so a solution will need to be devised if the two are ever to operate together. Hopefully an effort will be made to correct this since the facility of an AO system with a coronagraph and an IFU presents a unique combination of instrumentation, allowing new science studies to be pursued.

5.2.2 Upgrade to non-standard masks

The Lyot masks provided with OSCA have been found to be non-ideal when used with the NAOMI segmented mirror. The masks were designed without ever seeing an actual pupil plane image of the system, but instead from sizes of the telescope primary and secondary obtained from engineering drawings of the WHT tube assembly. Although the masking was provided at a level of 80% undersized primary and 120% oversized secondary, this was not large enough to mask the central four NAOMI segments. Since the central

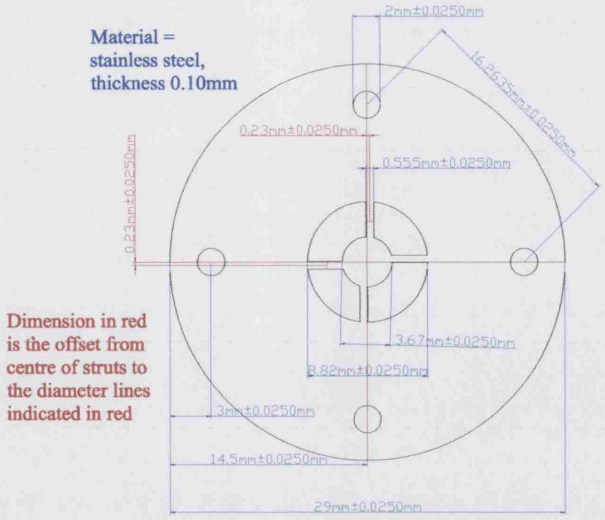

Figure 5.1: New lyot mask designed with a larger central obscuration to completely cover the central four inactive NAOMI segments. The vane masking is 50% bigger for increased error tolerances. The primary masking is at 90% to increase throughput given the other increased masking.

four segments are not active they will contribute to unwanted artefacts in the image if not properly covered. Figure 5.1 shows a new Lyot stop design with an enlarged central obscuration to resolve this.

The Lyot masks provided for NAOMI and OSCA in INGRID were also a non-ideal design. Although the primary and secondary masking was adequate, there are 3 secondary support struts at 120° intervals which are comparatively wide causing additional PSF artefacts. The struts should have been chosen to match with the diffraction pattern that the square segments of NAOMI already create, as shown in Figure 5.2.

The potential benefit of the Lyot mask illustrated in Figure 5.3 can be seen by looking at a pupil image such as that shown in Figure 4.23. In this image there is an obvious over-intensity around the NAOMI segments which could be reduced with the new Lyot mask. Hopefully this would improve contrast and the suppression of the object in the final image.

Much work has been done of late on the topic of pupil apodisation (e.g. Kasdin *et al.* (2003), Debes *et al.* (2002)), and various novel pupil plane masking techniques that provide high contrast regions in the final image. This is another area of research that could be

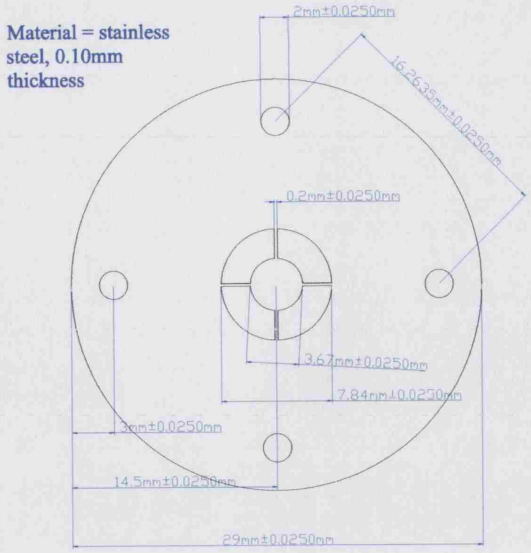

Figure 5.2: New lyot mask design, again with increased central obscuration to cover NAOMI segments. The four struts are for support only, this is for non-rotating Lyot masking. The primary masking is at 80%.

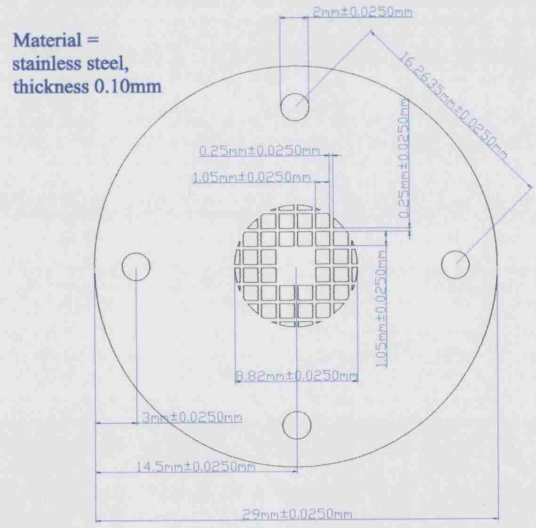

Figure 5.3: Novel Lyot mask design specifically designed for the NAOMI segmented mirror. Slight phasing errors between each segment cause an over-intensity about each segment post-occulting spot. This mask under sizes each segment by 80%.

investigated in the future and implemented at the WHT. The pupil designs could be placed in INGRID to provide cooled, high contrast pupil masking.

The pre-AO coronagraphic stop has not yet been implemented. The main reason for this was the difficulty of producing a 300μm circular dichroic deposit. The dichroic was to be made by Barr-Associates using a formula that was created for the large NAOMI dichroic. This transmits visible wavelength light to be used by the wavefront sensor but reflects all longer wavelengths. A coronagraphic stop made from this material would allow the star to be wavefront sensed in the visible but the science wavelength blocked, thus reducing scattered light from the DM and any other optics. It transpired, after many efforts by the manufacturer, that this stop was too difficult to create. However, the future plan is still to go ahead with the pre-AO stop (there is a space for it in the pre-NAOMI filter mechanism (Figure 4.37)). One idea is to use a thin gold deposit and apply photolithography techniques to create the 300μm spot. Gold is a good IR reflector and hopefully a thickness could be determined that lets through enough visible light to sense on while still reflecting a reasonable amount of the IR.

5.3 Other AO Optimised Instrumentation

The development of adaptive optics will continue to have a significant impact on the design of future instrumentation such as OSCA. Most AO systems are costly to build and implement and their running at a telescope requires considerable overheads in time (Marco *et al.* 2001) and staff training. It is therefore important to exploit the benefit gained using AO by including other instrumentation optimised to utilise it and extract the maximum amount of science. Some examples of instruments that benefit from an AO feed are:

- Spectroscopy - better centring stability of the source on the slit and higher throughput for improved spectral visibility. This includes integral field spectroscopy (such as OASIS) which uses optical fibre bundles to channel the light from different parts of an object to simultaneously take multiple spectra. With the use of AO the object is better resolved and thus if needed the fibres in the bundle can be positioned closer together to get individual spectra of finer, closer details of the object in question.

- Interferometry - adaptive optics is needed to improve the magnitude limit of these

systems by increasing the effective atmospheric coherence length r_0. By restoring the coherence of light over the individual apertures of the interferometer the sensitivity (fringe visibility) is greatly increased.

- CCD cameras - need to ensure that those used in conjunction with AO have a pixel size smaller than the diffraction limit of the telescope for a given wavelength range.
- Other novel coronagraph techniques using phase (Roddier & Roddier 1997) rather than amplitude modification such as the four-quadrant phase-mask put forward by Rouan *et al.* (2000). However this requires near-perfect centring, very stable wavefront correction and very high Strehl ratios which can only be produced by extremely efficient AO systems.

Adaptive optics also needs to be developed further so that good correction close to the diffraction limit can be obtained for shorter, visible wavelengths rather than the partial correction that can be achieved with the current near IR adaptive optics systems. Partial compensation occurs when the correction elements in the deformable mirror are appreciably larger than r_0. Although partial compensation is a useful improvement in image signal to noise when isolated point sources are being observed, for many closely spaced objects or extended objects the haloes of adjacent points may overlap, reducing image contrast. Thus in spectrographs where the power passing through the slit needs to be maximised or contamination from other sources avoided a higher degree of wavefront correction is required.

5.4 The future of coronagraphy

The focus for the next generation of high contrast imaging systems is mainly towards space-based systems. Since the major science drivers are the detection and characterisation of exo-planetary systems (both terrestrial and jovian type planets), the short fallings in AO bring into question whether these objects will ever be directly observable from a ground-based system. The detection level for a terrestrial planet in reflected visible light is 1 part in 10^9; for such requirements a space-based system would seem to be the only choice, especially in the near to mid future. We need only compare the contrast in Hubble images with those of the best AO systems (Figure 4.41) to appreciate this.

There are a number of mission proposals for space-based high contrast imaging, with the aim of direct observation of planets. Some of these are precursor missions to test new ideas and technology which will then be improved for more ambitious missions (e.g. the Jovian Planet Finder as a precursor for the Terrestrial Planet Finder). Others are survey missions to identify stars with potential planetary systems with subsequent plans to employ a suitable high contrast imaging technique to study the results. Summarised here are a few of these future missions and the technology they hope to employ to obtain the required suppression for imaging exo-planets.

1. The Jovian Planet Finder (JPF), (Clampin *et al.* 2001; Krist *et al.* 2003). This is planned to have a 1.5m aperture to detect jupiter-like planets in jupiter-like orbits, with a 10^9 contrast ratio. The telescope is planned to be an unobscured, off-axis design to improve contrast and throughput, while the optics will be super-smooth to reduce scattering to a minimum - essential for any high-contrast imaging system. It will use occulting masks of 0.8 and 1.6 arcseconds FWHM, of the gaussian-apodised design (the same type as the OSCA gaussian masks) and some form of shaped Lyot mask. A corrector optic and fast steering mirror maintain accurate positioning of the star behind the occulting mask. Speckles caused by optical errors can lead to false detections and make real detections difficult to identify. The JPF uses a modified field rotator which smears these speckles azimuthally but maintains the planet position. The background thus becomes smoothed to facilitate detections.

2. The Extra-Solar Planetary Imager (ESPI) (Lyon *et al.* 2003). This is another proposal for the design of a system capable of detecting Jovian-like planets. Unlike the previous design (JPF), this one utilises a 1.5m $\times$ 1.5m apodised *square* telescope. This allows detections of very faint sources to within 0.32 arcseconds of bright stellar objects. ESPI will also be equipped with a spectroscope to help characterise any detections. For some very close stars, ESPI will search for terrestrial-type planets. This design does not utilise a focal-plane occulting spot.

3. The Umbral Missions Blocking Radiating Astronomical Sources (UMBRAS) (Schultz *et al.* 2003). This consists of a 1m apodised space telescope with a separate free-flying occulter, 1000 to 15000km from the telescope. The benefits of such a system is that it significantly reduces scattered light since less light actually enters the telescope. Contrast ratios of 10^8 or greater have been simulated with this design.

4. Darwin (Ollivier *et al.* 2001). This is the European effort (ESA) to build a system capable of imaging terrestrial planets (in competition with NASA's TPF). To obtain the required resolution, Darwin plans to utilise six, 1.5m, formation flying Cassegrain-type telescopes acting as an interferometer. The system will operate in the mid-infrared where the contrast requirements are about 1000 times less ($\sim 10^6$) than those in the visible. The telescopes will be positioned at L2 to allow observations to go on uninterrupted by the usual orbital constraints. A spectroscope will be available to help determine the nature of any planetary atmosphere.

5. The Terrestrial Planet Finder (TPF) (Beichman 1998). There are a number of plans for the TPF: the two main categories are a visible/IR coronagraph or an interferometer operating in the infrared. To overcome problems of excess light from the local zodiacal dust, TPF plans to use larger telescopes which will provide the required sensitivity at a 1AU orbit (compared to smaller telescopes in more distant orbits). Various optimal apodisations have been modelled for use with the TPF (Gonsalves & Nisenson 2003) in an effort to find transmission functions that can provide contrast ratios of 10^{10} in certain regions close to a star.

Future plans to detect planetary systems from ground based telescopes are reliant on the development of extremely high-order adaptive optics systems on large telescopes (10-100m primary), with the aim of attaining Strehl ratios of 70-90%. This requires a corrector with a very large number of actuators and small inter-actuator spacing. The use of technologies such as MOEMS (micro-opto-electro mechanical systems) (Zamkotsian *et al.* 2003) could be suitable for such mirrors: obtaining a high mirror quality is the biggest challenge for such designs. For a planet finder type survey, numbers of the order of 10^4 have been quoted as the required number of actuators over an 8-metre aperture. The control system and computing power alone make this a considerable challenge.

Given that the current, much simpler AO systems frequently encounter various technical difficulties which degrade their performance from that predicted at their design stage, we must be cautious about the predictions for these future, complex designs and question whether they will be able to deliver the expected astronomy goals that are driving their creation. Computer simulations modelling new systems must be carefully and rigorously tested to ensure they are realistic; current systems can help as a basis for this. The huge technological leap required to go from the high-contrast imaging systems in use today to

the next generation systems is also a concern. Most of the space-based designs are natural progressions from previous planned designs (JPF → TPF) and the technological challenges do not seem as difficult for similar scientific discovery aims. It would seem, therefore, that the future in high contrast imaging with the best chance towards terrestrial planet finding is in the development of getting *large* telescopes into space.

Appendix A

Additional Photographs

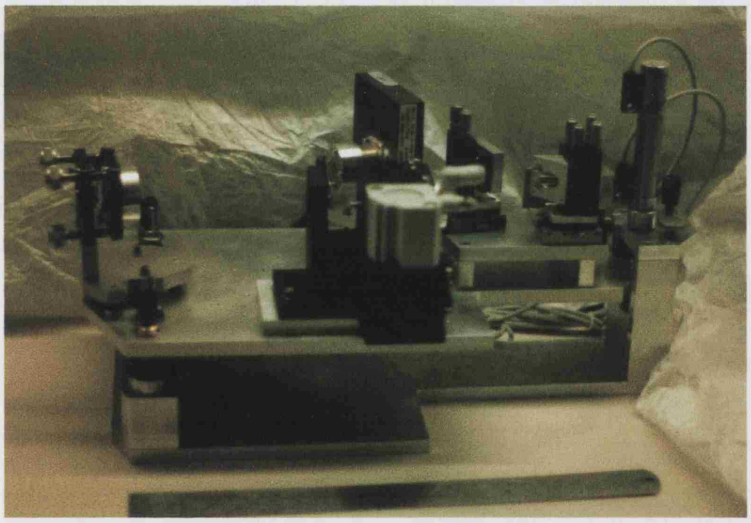

Figure A.1: First assembly of OSCA in the Optical Science Laboratory at UCL. Except for the focal-wheel assembly, OSCA is in the bare metal stage, prior to black anodising. A 30cm ruler is shown in the foreground for an indication of scale.

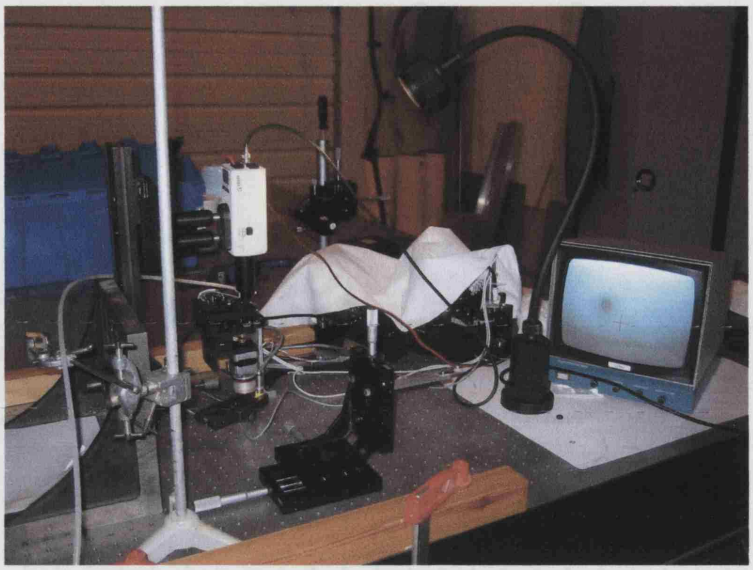

Figure A.2: Installing the occulting masks. The substrates are aligned using an xyz-stage micrometer adjustment. A microscope objective screwed into the video camera relays a magnified image to a TV screen. The centre of the reference mask has been marked on the TV screen.

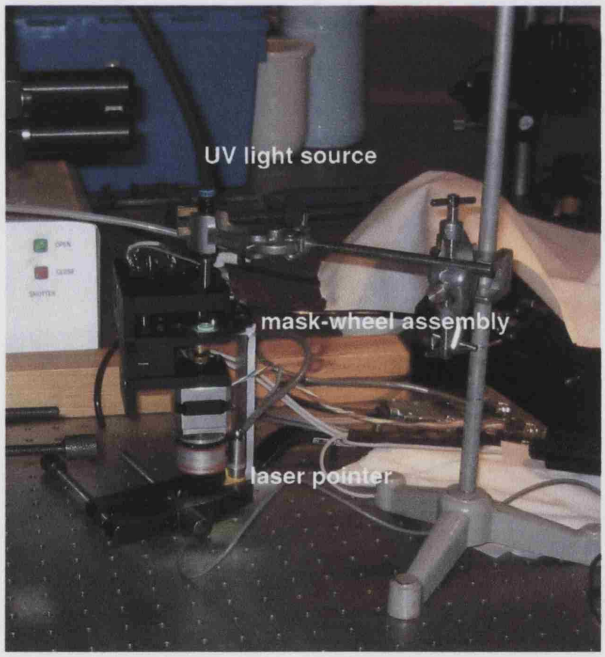

Figure A.3: Fixing the focal masks in place using UV-glue

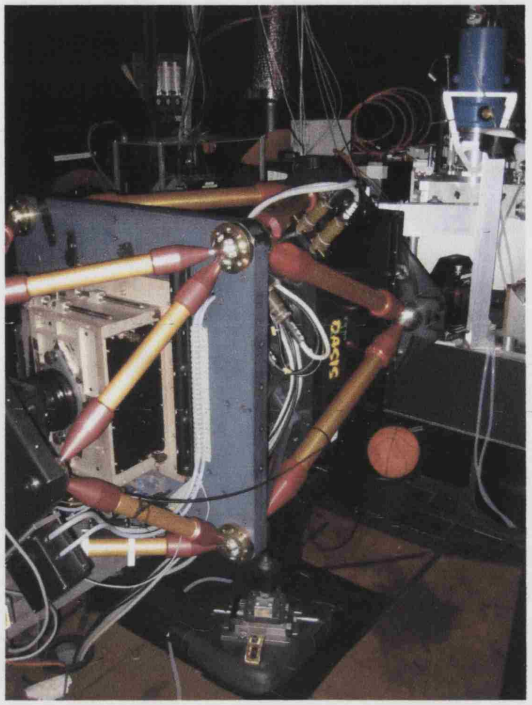

Figure A.4: OASIS is an integral field unit recently installed at GRACE at the WHT. When fully commissioned it will be available for use with NAOMI and possibly OSCA as well.

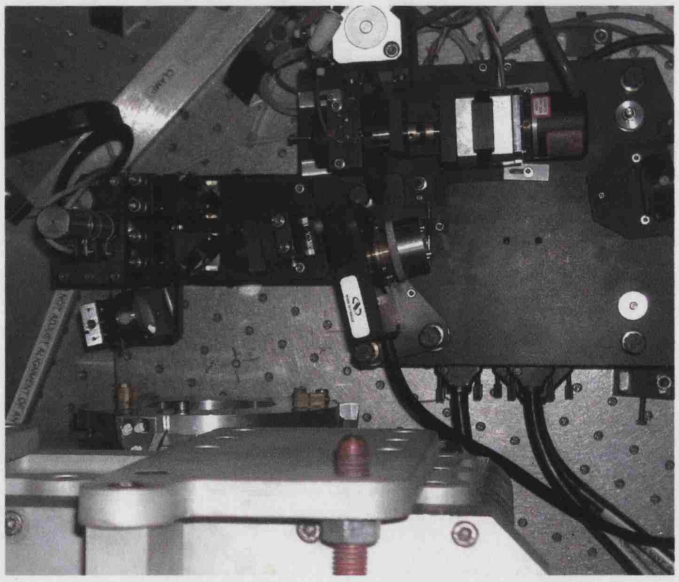

Figure A.5: Birds-eye view of OSCA

Figure A.6: The author avoids the obstacles to align the OSCA components, such as the Lyot stop and tilt of the focal plane mask assembly.

Appendix B

CAD drawings

All designs here were designed/drawn by the author unless otherwise stated.

B.1 OSCA Mirror Mountings

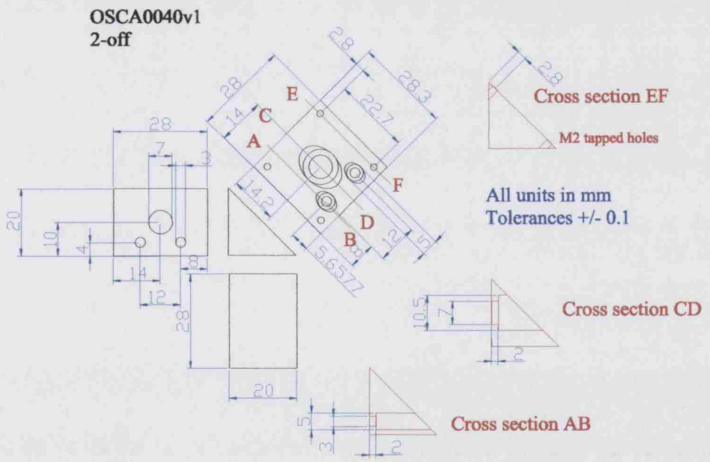

Figure B.1: Right-angled wedge mountings for flat-mirrors 1 and 3

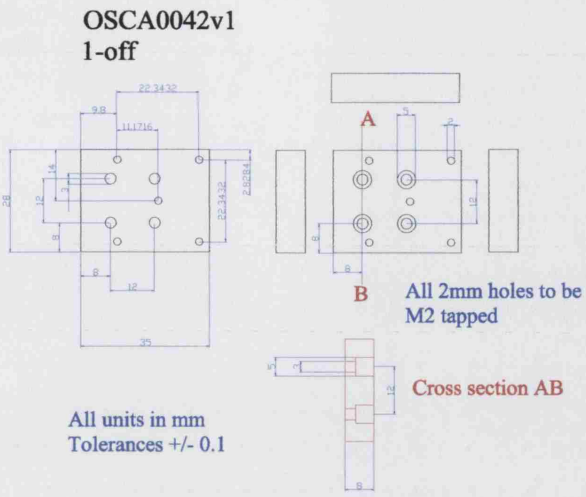

Figure B.2: Block mounting for flat-mirror 2

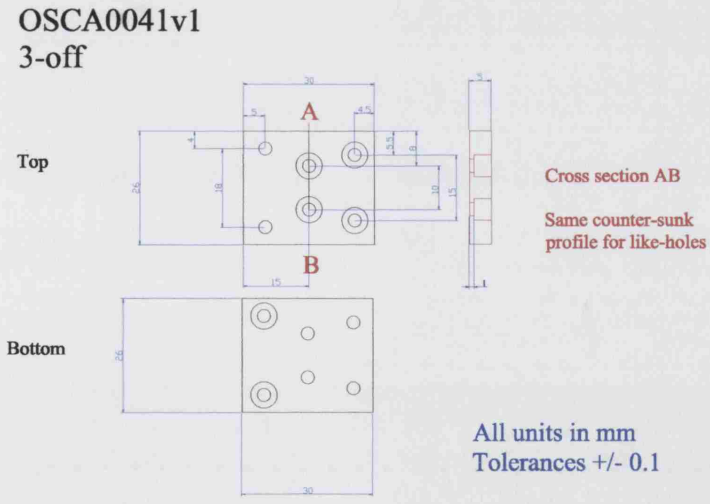

Figure B.3: Bottom spacers for flat-mirrors 1, 2 and 3 to give correct vertical placement

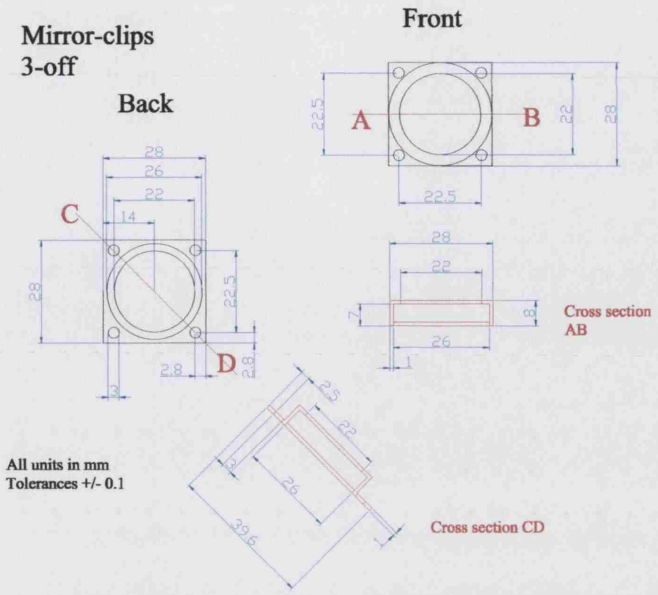

Figure B.4: Clips to attach flat-mirrors to the mountings (Figures B.1 and B.2)

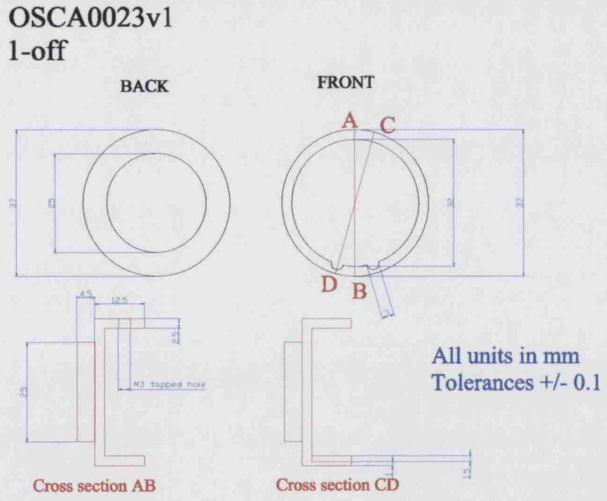

Figure B.5: Mirror holder for OAP1

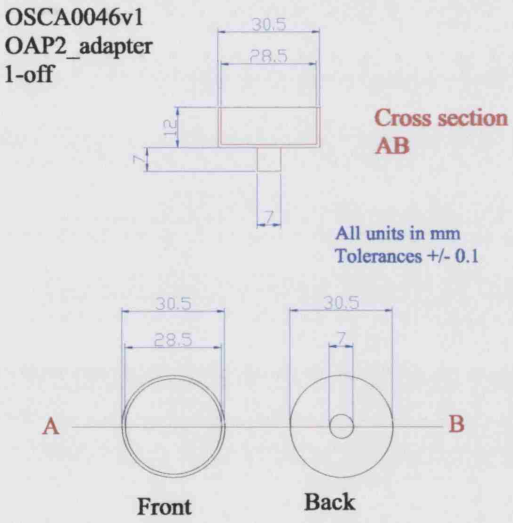

Figure B.6: Mirror holder for OAP2

B.2 Occulting mask wheel assembly

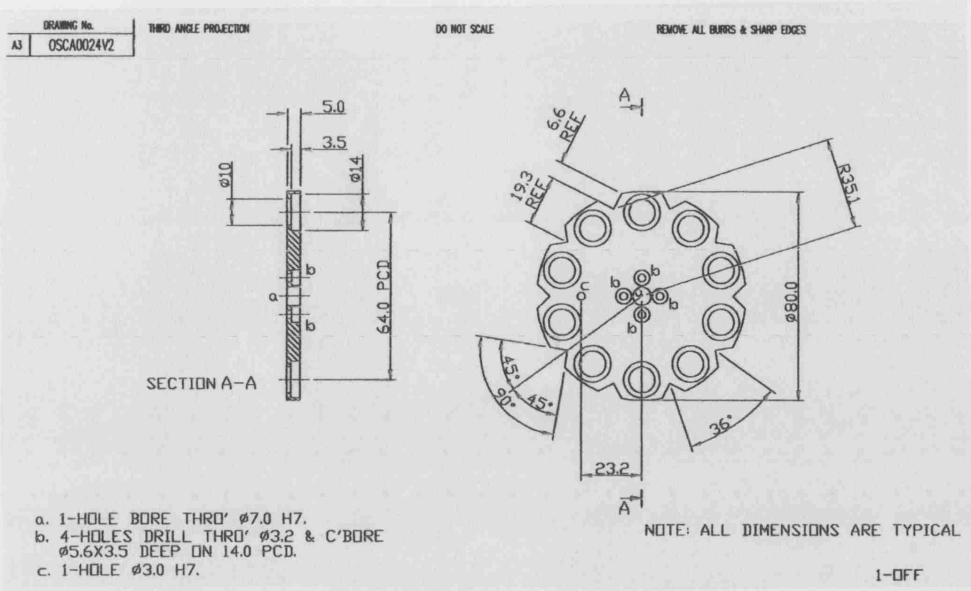

Figure B.7: Occulting mask wheel (Andy Charalambous).

BIBLIOGRAPHY

Aime, C., Soummer, R. & Ferrari, A., 2001, *A & A*, **379**, 697

ATC, 1997, *Technical Description of NAOMI*, Internal doc number aow/gen/ajl/7.3/07/97/technical description of naomi, ATC and Durham University

Baudoz, P., Rabbia, Y. & Gay, J., 2000, *A & AS*, **141**, 319

Beichman, C. A., 1998, *SPIE*, **3350**, 719

Boccaletti, A., Augereau, J. C., Marchis, F. & Hahn, J., 2003, *ApJ*, **585**, 494

Burrows, C. J., Krist, J. E. & Stapelfeldt, K. R., 1995, *AAS*, **27**, 1329

Carson, J. C., Eikenberry, S. S., Brandl, B. B., Wilson, J. C. & Hayward, T. L., 2003, *ApJ preprint*, **10.1086/376514**

Clampin, M., Ford, H. C., Illingworth, G., Petro, L. & Team, J. S., 2001, *AAS*, **33**, 1356

Debes, J. H., Ge, J. & Chakraborty, A., 2002, *ApJ*, **572**, L165

Doel, A. P. & Myers, R., 1999, *A Coronagraphic imager for the NAOMI Adaptive Optics System*, ING proposal, University College London and Durham University

Doel, A. P., Thompson, S. J. & Bevil, C., 2002, *OSCA Observers Guide*, OSCA manuals, University College London and Isaac Newton Group

Ellerbroek, B. L., Rigaut, F. J., Bauman, B. J., Boyer, C., Browne, S. L., Buchroeder, R. A. & 16-coauthors, 2003, *SPIE*, **4839**, 55

Golimowski, D. A., Durrance, S. T. & Clampin, M., 1993, *ApJ*, **411**, L41

Golimowski, D. A., Nakajima, T., Kulkarni, S. R. & Oppenheimer, B. R., 1995, *ApJ*, **444**, L101

Gonsalves, R. & Nisenson, P., 2003, *PASP*, **115**, 706

Goodman, J. W., 1996, *Introduction to Fourier Optics* (McGraw-Hill)

Hardy, J. W., 1998, *Adaptive Optics for Astronomical Telescopes* (Oxford University Press Inc)

Heap, S. R., Lindler, D. J., Lanz, T. M., Cornett, R. H. & Hubeny, I., 2000, *ApJ*, **539**, 435

Kasdin, N. J., Vanderbei, R. J., Spergel, D. N. & Littman, M. G., 2003, *ApJ*, **582**, 1147

Knude, J., 1992, *A & AS*, **92**, 841

Krist, J., Clampin, M., Petro, L., Woodruff, R., Ford, H., Illingworth, G. & Ftaclas, C., 2003, *SPIE*, **4860**, 288

Lane, R. G., Glindemann, A. & Dainty, J. C., 1992, *Waves in Random Media (IOP)*, **2**, 209

Lyon, R. J., Gezari, D. Y., Melnick, G. J., Nisenson, P., Papaliolios, C. D., Ridgway, S. T., Friedman, E. J., Harwit, M. & Graf, P., 2003, *SPIE*, **4860**, 45

Lyot, M. B., 1939, *MNRAS*, **99**, 580

Malbet, F., 1996, *A & AS*, **115**, 161

Marchetti, E., Brast, R., Delabre, B., Donaldson, R., Fedrigo, E., Frank, C., Hubin, N. N. & 13-coauthors, 2004, *SPIE*, **5490**, 236

Marco, O., Ageorges, N. & Sterzik, M., 2001, *PASP*, **113**, 397

Mouillet, D., Lagrange, A. M., Augereau, J. C. & Menard, F., 2001, *A & A*, **372**, L61

Myers, R. M., Longmore, A. J., Benn, C. R., Buscher, D. F., Clark, P., Dipper, N. A., Doble, N., Doel, A. P., Dunlop, C. N. & 13-coauthors, 2003, *SPIE*, **4839**, 647

Nakajima, T., 1994, *ApJ*, **425**, 348

Nakajima, T., Durrance, S. T., Golimowski, D. A. & Kulkarni, S. R., 1994, *ApJ*, **428**, 797

Nisenson, P. & Papaliolios, C., 2001, *ApJ*, **548**, L201

O'Hara, T. B., Meixner, M., Speck, A. K., Ueta, T. & Bobrowsky, M., 2003, *ApJ*, **598**, 1264

Ollivier, M., Mariotti, J. M., Leger, A., Sekulic, P., Brunaud, J. & Michel, G., 2001, *A & A*, **370**, 1128

Pasquali, A., Nota, A., Smith, L. J., Akiyama, S., Messineo, M. & Clampin, M., 2002, *ApJ*, **124**, 1625

Rigaut, F., Kern, P., Lena, P., Rousset, G., Fontanella, J. C., Gaffard, J. P. & Merkle, F., 1991, *A & A*, **250**, 280

Roddier, F. & Roddier, C., 1997, *PASP*, **109**, 815

Rouan, D., Riaud, P., Boccaletti, A., Clenet, Y. & Labeyrie, A., 2000, *PASP*, **112**, 1479

Schultz, A. B., Jordan, I. J., Kochte, M., Fraquelli, D. A., Bruhweiler, F. & 14-coauthors, 2003, *SPIE*, **4860**, 54

Smith, B. A. & Terrile, R. J., 1984, *Science*, **226**, 1421

Testi, L., Palla, F. & Natta, A., 1998, *A & A*, **133**, 81

Trilling, D. E., Brown, R. H. & Rivkin, A. S., 2000, *ApJ*, **529**, 499

Tyson, R. K., 1993, *Principles of Adaptive Optics* (Academic Press)

Wachter, S., Hoard, D. W., Hansen, K., Wilcox, R., Taylor, H. & Finkelstein, S., 2003, *ApJ*, **586**, 1356

Watson, S. M., Mills, J. P., Gaiser, S. L. & Diner, D. L., 1991, *Applied Optics*, **30**, 3253

Weinberger, A. J., Becklin, E. E., Schneider, G., Smith, B. A., Lowrance, P. J., Silverstone, M. D., Zuckerman, B. & Terrile, R. J., 1999, *ApJ*, **525**, L53

Wilson, R. G., 1995, *Fourier Series and Optical Transform Techniques in Contemporary Optics* (Wiley)

Zamkotsian, F., Camon, H., Fabre, N., Conedera, V. & Moreaux, G., 2003, *SPIE*, **4839**, 711